

Second chance for nuclear power?

Last week's conference on Chernobyl was a landmark in the development of nuclear technology, not least because of Soviet openness: was it a new beginning or the beginning of the end?

How do you turn a calamity into a celebration? The religious would advocate expiation or something of the sort. This is much what seems to have happened at Vienna last week, when a Soviet delegation unsure of the reception it would be given decided to offer an almost full account of the world's worst peaceful nuclear catastrophe. That this openness accords with what Mr Mikhail Gorbachev has been saying should not diminish the importance of what happened. Academician A. V. Legasov eloquently explained (see pp 3 and 25) that it is easier to be frank about triumphs than about disasters. But frankness paid handsome dividends. Chastened by Soviet candour, most listeners seem to have decided that more of the same would be jeopardized by hostile questioning. The result was a sense of camaraderie not witnessed in this field since the first heady days of the peaceful exploitation of atomic energy more than thirty years ago. Two things must be said about that development: first, nothing but good can come from closer collaboration on Chernobyl and on nuclear safety in general and from a stronger role for the International Atomic Energy Agency and, second, that the camaraderie by itself will not chase off the political problems of nuclear power.

The difficulty is that what went wrong at Chernobyl on 26 April could have happened anywhere. That is the plain truth, which no amount of technical comparison of different reactor types can possibly conceal. Moreover, there have been several occasions in the recent past when nuclear accidents, luckily smaller in scale, have been brought about because operators have chosen to disregard the regulations they are supposed to live by, or have been deserted by common sense and elementary caution. Who can forget the fire at the Brown's Ferry reactor of the Tennessee Valley Authority caused by a maintenance man's decision to use a naked candle flame in a confined space? The accident at Three Mile Island would probably not have happened if the reactor operators had followed the rules and not their hunches. The Canyon Diablo reactor in California, originally built with some of its components installed back to front, and then expensively modified, could have become an equally expensive monument to human folly if the errors had not been spotted.

Errors

The chain of events leading to the accident at Chernobyl stands out because of the sequence of errors the operators chose to make, all of which are consistent with what may have been the belief that a reactor operating at low power is inherently less dangerous than one operating at high power. Only some such conviction, coupled with the blend of complacency and arrogance that springs from a reputation for success, can explain why experienced people met the prospect that safety systems might frustrate a planned experiment by switching off those systems. Thus might a person with angina attempt to hide the pain in his chest by swallowing aspirin.

This is why no great purpose can be served by detailed comparison of the Soviet reactor with other machines elsewhere. It is nevertheless the case that the Soviet reactor embodied a recipe for positive feedback between reduced cooling efficiency and

increased heat production. So, as the cooling water turned to steam in the last few seconds before the reactor ran amok, the power production went on rising even faster. Reactor managers elsewhere may rightly claim that their machines do not have this particular fault, and that even a contrived approach to catastrophe would be more moderate. There is also much that may (and should) be done by means of electronic locks and keys to guard against contrived accidents such as that at Chernobyl. Intelligent (not bureaucratic) procedures are also valuable, as are the techniques worked out to help long-distance aircraft pilots, and others doing demanding but boring jobs, to remain alert. But nuclear safety engineers elsewhere cannot (and, to their credit do not) absolutely claim that "it couldn't have happened here in [fill in the blank]".

Procedures

How, with such a fuzzy task ahead of them, can the world's nuclear engineers hope to persuade their governments to continue building nuclear power stations? Their main argument is what it has always been, that for many countries there is no economic choice. But the frankness of the Soviet Union last week about the course and the physical consequences of the accident at Chernobyl also demonstrates that there are many countries that could not survive an accident on such a scale, economically and socially. So it is inevitable that governments will now take a keen interest not just in the way nuclear power stations are designed and built but also in the way they are run. The US Nuclear Regulatory Agency, which has a fearsome reputation among reactor operators for bumbling interference, is forever fining people on the grounds that they have broken agreed rules. Now, no doubt, similar systems of invigilation will be copied elsewhere. Nuclear engineers should not protest too loudly; in the long run, some such procedure may be a necessary, if transitional, means of restoring public confidence.

But even tougher regulatory regimes for nuclear energy will not always satisfy the public anxiety engendered by Chernobyl. Some governments will no doubt decide to follow Sweden (which, with one of the best safety records in the world, has decided to build no more nuclear power stations) or even Austria (which has decided to dismantle its own plant, which has not been switched on). There is no reason why particular governments should not take such apparently iconoclastic decisions, which can always be reversed when circumstances change. But in the major industrial economies without indigenous cheap fuel, the two big creditors on the international trade markets (Japan and West Germany) in particular, the economic case for perseverance cannot be dismissed. But after Chernobyl, even there a tighter regulatory regime will be an advantage.

The obvious snag in the regulatory approach to safety is that, in the presently polarized state of opinion on nuclear power in many Western countries, there is a kind of Gresham's law that drives safety standards downwards. Soviet spokesmen in Vienna last week were right to complain at the decisions of European governments to ban food imports from the East whose radioactive content was an order of magnitude more than necessary. The British government, which banned the sale of new lambs



because the caesium content of some of them exceeded another arbitrarily low limit, may like to know that the government of Argentina fixed an even lower limit to safeguard its people (and the customers for its meat exports) from Chernobyl fallout, secure in the knowledge that tropospheric air masses do not cross the Equator. This is why one of the most important tasks with which the international agency has been saddled is that of negotiating standards of radiological safety which are recognized internationally to be sensible. That will not be easy.

Regulation

Regulation is also a means by which anxiety about nuclear safety may be excited. Each enforcement of a regulation may be represented as a sign that catastrophe has been only narrowly averted. In Britain, for example, where the nuclear issue is well on the way to being as contentious as in the United States, with the recent practice of making public announcements of all radioactive discharges to the environment, each few becquerels that find their way into the atmosphere or some water-body are greeted as a sign of the ending of the world by those committed to the view that nuclear energy is by definition an abomination. But the temptation to abandon openness must be resisted. The best, indeed the only hope, is that people will learn from repetition that radioactivity does not differ from other environmental pollutants in being dependent for the damage they do on their amount. Chernobyl was a serious disaster (and could, with bad luck, have been a lot worse.) Most radiation scares are quite different, but may malevolently be used to stir up trouble.

That is only one reason why it must be hoped that the regulatory approach to nuclear safety will in due course be overtaken by a more positive solution of the problem. In at least one respect, the Soviet system appears to have contributed substantially to the successful handling of the emergency: the educational system makes Soviet citizens knowledgeable about a variety of technical matters, and better able to appreciate the stochastic character of radiation injuries than is likely to be common elsewhere. It is also true, of course, that the Soviet system requires an unnatural degree of compliance with centrally laid plans, such as the decision that young children should be evacuated separately from their parents from the 30-km zone around the reactor. The moral for other would-be civil nuclear powers is that there is much to be gained from a deeper general understanding not merely of the putative benefits of nuclear power, but also of the risks. Even more openness will be needed.

None of this will ensure that operators behave responsibly. One of the ingredients missing from last week's Soviet report, for obvious reasons, is a full account of the reasons why the plant operators thought it necessary to depart so far from normal practice at the damaged reactor. For that matter, what is known of the frequency of corner-cutting without mishap there and at other power stations? And what of the degree to which the Soviet system, now more than ever one in which liberal people tolerate only grudgingly state institutions that have lost general respect, may have contributed to the accident by engendering cynicism about even sensible rules and regulations?

The counterpart in the West is the willingness of contractors to skimp on the quality of equipment supplied or of managers to discard warnings from professional people in the pursuit of what would be called "norms" in the Soviet Union. The loss of the US space shuttle last January refers. The only remedy that will in the long run work is that professional engineers, whatever their status in a hierarchy, should have the right to speak their minds and to be listened to. If it is accepted that the general public should be well-informed, should not the professional people who carry the responsibility enjoy a degree of independence they are now denied? That, of course, will be a hard reform to implement, not merely in the Soviet Union. The Soviet account of Chernobyl is that of a thoughtless crew eager to finish off a tedious chore. Would they have behaved like that if they were in charge, not merely shift-workers? □

Chips in big boxes

The personal computer manufacturers are in calm waters, and may become sleepier.

CURIOUS things are going on at the bottom of the computer market, the battlefield on which dozens of companies have lost their shirts (or, more cannily, other people's) in the past three years. Surprisingly, prospects are looking up for the companies that have managed to survive, while even business (the prospects for the next quarter) is improving. What can have changed? And why are many manufacturers of electronic components, especially in the United States, still wailing loudly at what they assert must be unfair competition from elsewhere?

What has been happening is an illustration of a familiar and easily recognized economic phenomenon. There was a time, roughly a decade ago, when toy manufacturers were seeing what uses they could make of the semiconductor chips that much more august manufacturers were used to building into mainframe computers. Perversely, some toy-makers built the chips into toy computers, which the younger generation occasionally lent to its elders. Not much time was needed for entrepreneurial elders to sense a market for machines selling for hundreds, not thousands, of dollars, which is how the trade in personal computers sprang to life. But as with the British railway construction boom of the mid-nineteenth century, eager investors were much more numerous than those among them who could grab a substantial share of the market. On the railways, the outcome was predictable. Companies swallowed each other and became bigger in the process until there was nowhere else for them to go, whereupon the incentive for technical change melted away. The result is modern British Rail.

Computer users are more fortunate, so far at least. The shake-out has been traumatic for many companies. Outright bankruptcy has been more common than merger and amalgamation. But consolidation and ossification have not followed. Part of the benefit, for users if not for established manufacturers, is that there seems to be an endless supply of upstart companies willing to compete with the established fellows on price, but that are unable consistently to beat them on performance. Both kinds of companies have also benefited, the upstarts from the money they have sometimes made, the established fellows from the demonstration provided at little cost by the others that there is a huge and sustainable market to work. The competition has become so volatile that the big fellows fear they will be swallowed by the upstarts; some are even thinking they should compete on technology.

That is why the big fellows are at last waking up. IBM's personal computer (called PC) is now long in the tooth, but there is a better version on the way. It may be more significant that Digital Electronics has thought it worthwhile to sell a cheap (or cheapish) computer that is compatible with its own minicomputer called VAX. Even in Europe, once nearly disappeared companies (such as the British Acorn) are talking as if they have a future again (under the Olivetti umbrella).

So is the shakeout at an end? Unfortunately, at least for manufacturers, no. The plain truth is that the machines now on the market are still rudimentary devices, whose big boxes are still mostly filled with old-fashioned wiring. People talk of the time when it will be possible to carry the power of a mainframe computer in a shopping bag, but the lap-top computer has only just become practicable and affordable.

It will be interesting to see whether the new chip of which Intel is now boasting will be that much more capable than Motorola's new product at enabling the design of decisively superior machines, but past performance does not suggest that even the thrusting entrepreneurs in the business are that adventurous. Their products are still most of all conspicuous for the size of the boxes they inhabit. There is a long way to go before their promise is delivered. □

Chernobyl

Soviet frankness creates sense of solidarity

Vienna

THE chief and perhaps only beneficiary of the Chernobyl accident may be the International Atomic Energy Agency (IAEA), which organized last week's technical meeting on the subject. The meeting finished with no fewer than thirteen suggestions for extra work on reactor safety the agency should pursue.

But although the agency in the past few months has been able to increase its nuclear safety budget by \$2 million a year, or roughly a third, by reallocation of present resources, IAEA will have to wait for a governors' meeting a month from now to learn whether member governments will back last week's recommendations with extra funds. The obvious danger is that governments may not be as strongly committed to nuclear power as the reactor people and other specialists last week.

Meanwhile, it is generally accepted that the technical meeting was a huge success, chiefly because of the frankness of the Soviet report on the reactor accident. (For a more detailed report, see p. 25.) Thus the long catalogue of incomprehensible errors leading to the accident, which is a sufficient explanation of it, has not restrained the Soviets from listing the technical defects of their reactor design.

The meeting nevertheless began nervously. The leader of the Soviet delegation, Academician V. A. Legasov, deputy director of the Kurchatov Institute in Moscow, explained that at the outset he and his colleagues were uneasy "that we were sharing not our successful experi-

ence with nuclear ice-breakers or an encounter with Halley's comet, but a painful experience, which was a great tragedy for us and which has worried many people elsewhere in the world".

Legasov had introduced the Soviet report by asking for constructive suggestions and criticism, and finished by offering continuing cooperation on nuclear safety and related issues. Almost all other delegates set out consciously to create what Legasov called "a sense of solidarity". Delegates have argued that the Soviet delegation would have been less than forthcoming if subjected to outspoken criticism.

One result of that strategy may nevertheless be that some questions in which IAEA member governments are interested, such as the apparent delay in the provision of accurate information from the Soviet Union, were not fully explored.

Even so, more than 700 questions were submitted by fewer than half as many delegates. The Soviets will answer in writing those not dealt with last week.

Delegates, including those from the Soviet Union, have been most of all puzzled by the apparently capricious behaviour of the operators of the reactor on 25 and 26 April, for which there seems no convincing explanation. But some of the control-room errors, and especially the failure, early on the morning of 26 April, to reset the working power level for the automatic control-rod system, appeared familiar to some Western plant managers.

Measurements of the radioactivity released from the reactor have confirmed that the accident is the worst so far. Fallout in the Soviet Union is estimated at 50 megacuries, with as much activity again in gaseous form. Doses to some people within the 30-km evacuation zone are estimated to have ranged up to about 70 rem, and plans are being laid for a workshop at which the details of a follow-up study extending over decades can be discussed. But the meeting took the view that the Soviets had overestimated the total collective radiation dose to the population of Western Russia, perhaps by a factor of 10.

One striking feature of the Soviet report is its coolness towards bone marrow transplantation in the treatment of those exposed to large doses of radiation, some as large as 1,600 rem. The report says that extensive burns, many by beta-irradiation of the skin, excluded transplantation for the most seriously affected patients, while those in whom it was used included some in whom spontaneous recovery of marrow

function caused host-graft reactions. These, the report says, may have contributed to the deaths of two people.

Although fallout external to the Soviet Union has been postponed to future discussion, the responses of some Western governments were roundly criticized by Professor L. A. Ilyin, vice-president of the Soviet Academy of Medical Sciences and director of the Moscow Institute of Biophysics. He saw no need for the distribution of iodine tablets in Poland nor for restrictions on food imports from the East by West European governments.

John Maddox

Cuts hurt basic science in France

THE final figures for the 1987 French budget show that the minister in charge of research, physicist Alain Devaquet, has had some success in protecting scientific research from the ravages of a monetarist government. But François Gros, ex-director of the Institut Pasteur and science adviser to successive prime ministers in the previous socialist government, last week criticized current policies.

Gros has just returned from visits to Japan and the Soviet Union, which he sees as rising competitors for French biologists. He says that an apparent increase of 10 per cent in the budget of the Centre National de la Recherche Scientifique (CNRS) shows "a degree of protection for basic science" but should be seen against an 8 per cent cut imposed in May. The budget will now be 2 per cent less than it was to have been in 1986.

Up to 500 posts for technicians and a planned 2 per cent increase in research staff are also likely to be cut. Under the previous government, France had been increasing its competitiveness in science in relation to other countries, but that trend would probably now reverse, said Gros. "Increase in competitiveness comes from increase over inflation." The need for advanced equipment and technicians is growing rapidly in biology and real increases are necessary to stay ahead.

Although the basic research budget is to be held at 1985 levels, ministry figures confirm cuts in support for applied research. According to Gros, those receiving funding under the previous government's "mobilisation programme" for biotechnology, for example, have suffered setbacks of nearly 50 per cent.

The ministry of research, however, says the planned spending levels of the previous government were "unrealistic", and the applied programmes ill-planned and wasteful. Relevant programme agencies should be forced to be more practical by earning their own way, raising contracts with industry and selling services, the government argues.

Robert Walgate

Plan for action

The several technical sessions of last week's meeting selected the following topics for research, study and consultation under the aegis of IAEA.

- Further study, experimental as well as theoretical, of the accident.
- Design of the man/reactor interface.
- Training (and perhaps accreditation) of operators.
- Promulgation of agreed safety standards.
- Firefighting at nuclear plants.
- Chernobyl fallout.
- Decontamination techniques.
- Exchange of uniform monitoring data.
- Estimation of biological effects (especially for the 135,000 evacuees).
- Long-delayed effects.
- Acute radiation injuries.
- KI side-effects.
- Dosimeter inadequacy.

West German research

Industry's subsidy attacked

Hamburg

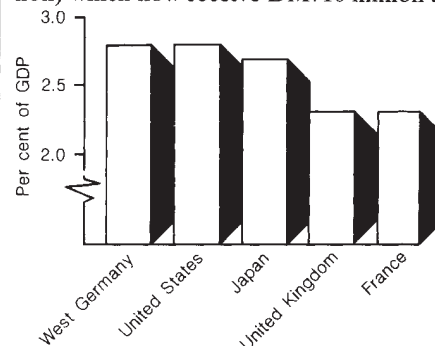
WITH a healthy balance of payments, a strong currency and low inflation, West Germany's economy is the strongest in Europe — and much of that strength is based upon successes in high-technology industries. Despite all this, there is concern that West Germany might not be keeping up with other countries in technology and scientific knowledge. And there is controversy over just what to do about it.

Decisions made by BMFT, the Federal Ministry of Research and Technology, hold the key to many of the major research and development projects, especially those requiring long-term investment. The minister at the head of BMFT, Heinz Riesenhuber, has his own ideas on how best to compete with the United States and Japan in fields such as biotechnology, space and information technology — he watches the markets to try to identify potentially profitable areas at an early stage and then sets up subsidy programmes to help industry to develop competitive research independent of government.

This "private enterprise" approach is

very much in line with the policies of Riesenhuber's party — he is a Social Democrat — and it seems to be paying political dividends, as he is regarded as a leading light in Chancellor Kohl's cabinet. The scientific value of BMFT's decisions is disputed by some experts, however. The nuclear power programme is one that has been criticized: more than DM3,000 million has been sunk into work on the fast-breeder reactor over the past decade or so, while high-temperature reactor technology has been neglected (only one prototype reactor has been built, in Hamm) and support for alternative energies such as solar power has been paltry. But Riesenhuber's ability to change priorities for BMFT's spending is severely limited: of a total civil budget of DM18,600 million, BMFT is responsible for DM7,500 million, but nearly three-quarters of that is tied up in long-term projects, support for major institutions and West Germany's contributions to international projects such as CERN and the European Space Agency. What should be done with the remaining funds — DM2,502 million last year — is hotly disputed.

In 1985 the nuclear programme took a large share of BMFT's industrial subsidy — DM740 million — leaving DM1,780 million with which to influence research trends in the rest of industry. Riesenhuber has favoured smaller companies (with an annual turnover of less than DM500 million) which now receive DM710 million a



West Germany's total research spending is 2.8 per cent of gross domestic product, which compares favourably to that in Japan and the United States. Two-thirds of this is met by trade and industry and one third by government.

year, whereas the larger companies now get around DM856 million compared with DM1,204 million in 1981.

This government support has to be matched with an equal investment from the company receiving the subsidy, but this has not prevented the accusation that government money is being spent extravagantly to support research that the companies would otherwise have to pay for themselves to survive in a free enterprise economy.

In particular, the Christian Democrats' coalition partners, the Free Democrats, are strongly opposed to the plan to subsidize Siemens and Philips to the tune of DM320 million over the next few years for the development of the four-megabit chip.

Riesenhuber, however, remains convinced that BMFT is following the right path. Support of DM1,000 million to 1989 in biotechnology is intended to achieve "jumps of perception" (see *Nature* 316, 287; 1985) and the number of scientists carrying out basic research in information technology should rise to 4,000 in 1992 from today's 1,700. And despite the problems of struggling with the very high cost of space research, Riesenhuber hopes to be able to support the Hermes shuttle project as well as participating in the US space station and the Ariane-5 launcher.

On nuclear power and alternative energy research, events and political expedience are causing some rethinking of BMFT policy. In the aftermath of Chernobyl, the views of environmental groups and the Federation of Solar Energy, an association of institutions and companies working on wind and solar energy, appear more attractive and Riesenhuber has announced that his ministry will support every promising project dealing with wind and solar energy.

Jürgen Neffe

French science

Researchers all at loggerheads

MEMBERS of the principal French researchers' trades union, the Syndicat National des Chercheurs Scientifiques (SNCS), earlier this month halted government plans to increase the power of universities over the Centre National de la Recherche Scientifique (CNRS), which funds the cream of French science, by walking out of a key decision-making meeting.

This action, according to SNCS officials, forced the new management at CNRS to negotiate with them over the issue, which it had refused to do earlier. The conflict centres on the reconstitution of the Comité Nationale, the recently dissolved collection of 44 elected specialist panels that makes the final decisions on appointments, promotions and grants at CNRS. The previous Comité was disbanded after it was declared illegal, leaving newly appointed scientists unsure whether they have posts at CNRS or not.

The proposition that university teachers and professors should have increased representation at the expense of CNRS scientists has provoked bitter opposition from SNCS, containing as it does a large proportion of CNRS scientists.

According to an SNCS official, the CNRS management is independently opposing the changes, and conflicting proposals from both management and SNCS are now going forward for the

ministry of research and higher education to disentangle. The union objects to two major proposals. The first is that Comité members would be from a single list, mixing both full-time CNRS scientists and university staff. In some disciplines, CNRS staff are completely outnumbered by those from university, so this would represent a total loss of power of CNRS researchers. The second proposal is that the number of middle-ranking researchers and technicians in each Comité panel would be reduced. The net result, the union would say, is that power over the CNRS would be too concentrated in the hands of the "mandarins".

The real issue is the long-standing competition between CNRS and the universities. Some university professors and departments, particularly in the humanities, social sciences, law and medicine, remain extremely reluctant to accept the power of CNRS, which has remained partly separate from the university system, despite the existence of many joint laboratories. Some objectors still go so far as to demand the dissolution of CNRS and its medical counterpart, INSERM, and a private bill (not sponsored by the government) to exactly that effect is to be read in the French National Assembly next session. The bill is unlikely to become law, but is certain to sustain conflict.

Robert Walgate

Ecological warning

Threat of floods from saline lake

THE resort town of Muysaly in the Soviet Union's Kazakh republic is in imminent danger of being "wiped off the face of the Earth", according to an inspector from the Soviet water resources ministry. The threat of a major ecological disaster comes from the saline Balkuduk lake, which has received no less than 80 million cubic metres of waste water from the industrial area around the town of Pavlodar.

Speaking on Moscow radio, the ministry inspector Vladimir Denisov warned that the lake is now so overburdened with water that a heavy storm might be enough to release flood waters that would not only engulf Muysaly, but also cause severe problems of pollution, with the polluted water being drained in to the river Irtysh, the primary water source for Kazakhstan and West Siberia. The pollution would cause millions of roubles of damage which would take several years to repair.

The management of a tractor plant, the biggest user of the Irtysh water (10,000 m³ daily), pays no attention to the conservation problems said Denisov; nor does the Ministry of the Chemical Industry, which has a major responsibility for industry in the area. The officials who should be most concerned with pollution, Denisov said, put their efforts instead into telephoning Moscow to obtain a stay of implementation of anti-pollution measures. Directives on building dykes and installing anti-pollution equipment are being ignored.

Water supplies to Kazakhstan, and the whole of Soviet Central Asia, have recently become a sensitive political issue. Russian writers of the "Villager" school have condemned plans to divert the north-flowing rivers to irrigate the arid steppes on the grounds that it would entail the destruction of treasures of Russian archaeology. The peoples of Central Asia, on the other hand, see such criticisms as a Russian move to establish superiority over the other ethnic groups of the Soviet Union. The recent Politburo announcement that it has been deemed "expedient" to end work on the diversion scheme "in connection with the need for further study of the ecological and economic aspects of the problem" will undoubtedly increase this resentment.

Denisov's revelations, therefore, however firmly based, are likely to be treated with suspicion by the inhabitants of Kazakhstan. His attacks on the Pavlodar local authorities will almost certainly be interpreted by the locals as a ploy to suggest that, as the Kazakhs cannot manage their own water, they have no right to extra supplies from the heartland of Russia.

Vera Rich

Biological weapons

New view from the Pentagon

Washington

THE new possibilities offered by recombinant DNA technology seem to have led the United States to change its view of the threat posed by biological weapons. US representatives at the second review conference of the 1972 Biological Weapons Convention, which starts this week in Geneva, will argue that it is now possible to produce effective biological weapons quickly and clandestinely. The key problem of preventing them from affecting friendly forces seems to be surmountable.

Because the United States believes the convention cannot be made verifiable, it is likely to resist diplomatic initiatives to strengthen it. US officials stress that the United States has no intention of abrogating the convention, and that it opposes the production of biological weapons.

But it is likely to resist expected proposals to clarify key definitions, such as the

distinction between offensive and defensive research, on the grounds that to do so would probably lead to a two-tier regime. Officials say the United States will also raise again its charges that the Soviet Union has repeatedly violated the terms of the convention by using toxin weapons in Afghanistan and South-East Asia.

The convention was drafted in 1972 and has since been signed by 102 countries. It forbids signatories from producing, stockpiling or using offensive biological weapons, but allows defensive biological weapons research. Critics contend that much ostensibly defensive research — for example, into vaccines against biological weapons — could also help in the manufacture of more fearsome offensive weapons. The Pentagon estimates that the size of the defensive biological weapons programme will increase from \$31 million in 1984 to \$63 million in 1987.

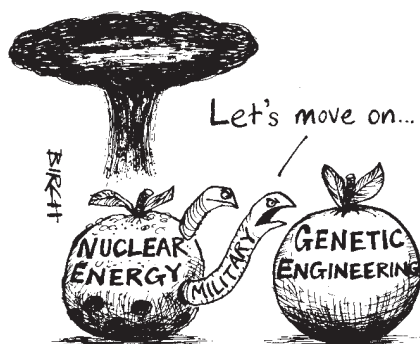
The administration's changed view of the feasibility of developing biological weapons was spelled out recently in testimony by the Department of Defense to the House of Representatives' intelligence committee. The testimony is unusual in that it was not requested by the committee, but volunteered by the department. Douglas J. Feith, deputy assistant secretary of defence for negotiations policy, wrote that "new technology has exploded the standard ideas about biological weapons that prevailed ten or more years ago".

Feith says that although biological weapons had previously been considered of questionable value because of difficulties of control and storage, it is "now possible to synthesize biological weapons agents tailored to military specifications" by, for example, "circumventing immunogens or antigens that the other side is suspected to possess". Together with advances in scale-up technology, this means that "the prevailing judgement of years ago that biological weapons are not militarily sustainable is quite unsustainable". While some independent experts are sceptical about quite how easy it would be to produce an effective engineered weapon, the Pentagon is not in a mood to make changes.

Feith says that biological weapons favour offence over defence, because it is easy to produce new agents but difficult to develop antidotes. He also claims that the Soviet Union has built a "large organization to develop and produce offensive biological weapons", and repeated the US claim that the Soviet Union has used mycotoxin weapons banned by the convention in South-East Asia. Although recent data question this claim (see *Nature* 321, 554; 1986), Feith says the fact that "mycotoxins occur in nature in certain colder areas ... has made it easy for states to refuse to come to the unpleasant conclusion that biological weapons have been used and that something should be done about it". Because of the controversy about "yellow rain", the Soviet Union "can hardly have failed to observe that the costs of biological weapons use have proven manageable, indeed virtually nonexistent".

US research into biological weapons has been hampered by a permanent injunction granted two years ago to Jeremy Rifkin's campaigning organization, the Foundation on Economic Trends, which prevents the Army from going ahead with its plans to build a \$300 million biological aerosol test facility at Dugway Proving Ground in Utah until an environmental impact statement has been prepared. Rifkin has filed suit again this week, this time seeking to halt the Army's entire biological weapons research programme pending environment impact statements. Rifkin has also established a \$100,000 whistle-blower's defence fund for microbiologists harassed or losing employment for publicizing illegal biological weapons research.

Tim Beardsley



Energy research

Tapping the tropical seas

Tokyo

JAPAN's industry and electric power companies are set to develop a vast pool of untapped energy — the tropical Pacific Ocean. This month they will help to launch a non-profit fund for the development of ocean thermal energy conversion (OTEC), aimed at helping the islands of the South Pacific.

OTEC is simple in principle. All you need is ready access to warm water from the ocean surface and cold water from its depths, resources that Pacific islands have in limitless amounts. In a closed system, warm ocean surface water (about 25°C) is pumped into heat exchangers to vaporize ammonia or freon which expands to drive a turbine. At the same time, cold water (about 4°C) from the deeper ocean (600–1,000 m) is pumped up to condense the vapour in separate heat exchangers, allowing the cycle to start again.

One of Japan's leading researchers in the field is Professor Haruo Uehara of Saga University in Kyushu, who in collaboration with industry has established two experimental plants; a 50–60 kW plant in Tokunashima island south of Kyushu, built by the Kyushu Electric Power Company, which operated from 1983 to 1985; and a 75 kW plant at Imari City near Saga University, built at a cost of ¥300 million (£1.3 million) with funds from the Ministry of Education, Science and Culture.

Uehara's next step is to design a commercial plant of 3–10 MW in collaboration with Nippon Kokan, a giant steel producer that is branching out into new fields. But even such small commercial plants will cost thousands of millions of yen to build and will be beyond the reach of their best potential customers, the islands of the South Pacific. This is where the soon-to-

be-established fund comes in.

More than 100 Japanese companies including the Tokyo and Kyushu Electric Power companies, Nippon Kokan and Mitsubishi Heavy Industries are expected to participate as supporting members of the fund which, with an initial operating budget of ¥500 million (£2 million), will provide grants to developing nations wishing to set up OTEC plants. The main backer, however, is expected to be the Japanese government.

One of the most likely first customers for an OTEC plant is the Republic of Nauru, a tiny Pacific island on the equator, which already has experience of operating a 100-kW experimental plant.

Several different designs are envisaged, including offshore floating plants, underwater shelf-based plants, and plants based on land. And they could include a marine hotel, aquarium and desalination plant.

But will the vast amounts of cold water pumped out of OTEC plants upset the marine ecosystem of the islands? Even a 10-MW plant requires a flow rate of about 100,000 tonnes per hour. Not so, say advocates of OTEC. In fact, they claim the cold nutrient-rich water is positively beneficial and can be used for the cultivation of abalone and fish.

The cost of OTEC-produced electricity for a 3-MW plant is a hefty ¥50/kWh, because of the high capital cost of the plant but for a 100-MW plant this figure drops to ¥19, according to Uehara, and OTEC would then be capable of competing with other sources of electricity in Japan. The most likely first candidates for OTEC power in Japan are the Okinawa and Izu islands, both of which are bathed by the warm waters of the Kuroshio current and lie close to cold deep water.

David Swinbanks

Archaeology

Conflict goes on over Congress

Southampton

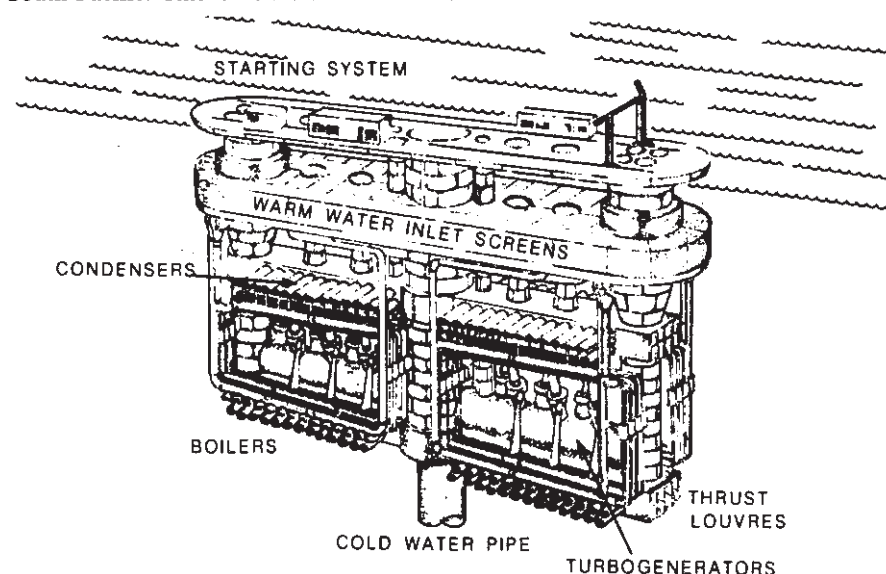
WITH the start last Monday of the troubled World Archaeological Congress here, the organizers began to count the cost of their ban on scientists from South Africa and Namibia and the subsequent removal of recognition of the meeting by the International Union of Prehistoric and Protohistoric Sciences (IUPPS — see *Nature* 319, 251; 1986). Clearly the end of the congress will not be the end of the story.

In spite of the problems, about 1,200 people turned up to register, which was fewer than the best hope of 2,000, but the loss of about 420 people who withdrew after the ban was enforced has hit particularly hard because these were mostly well-funded scientists from the United States and Western Europe. The total budget of the congress, estimated at about £0.5 million, has suffered further from the withdrawal or reduction of sponsorship, particularly from the United States. In consequence, administrative support has been cut back, and the number of travel grants for participants from developing countries has been reduced.

The congress has been particularly vulnerable to financial issues because of the novel features the organizers claim for it. The emphasis has turned away from Europe and North America and centres on particular themes to which archaeologists from developing countries can contribute more easily. The organizers have also invited several people recognized by their communities as being experts on their own cultural heritage. The bias towards participants from developing countries has required a greater reliance than usual on external support.

This shift in emphasis, however, is seen as extremely important by the organizers, headed by Professor Michael Day (St Thomas's Hospital Medical School, London), and has heightened the conflict between them and IUPPS, which is seen by many to look mostly towards European issues. This conflict will be the key issue at a plenary session to be held on Saturday 6 September, and it is hoped that the differences that have arisen over the ban on South African and Namibian scientists can be resolved before the official IUPPS World Congress scheduled for Mainz in 1987 takes place. If they cannot agree, Day believes that further "unofficial" congresses may result. But it is the hope of Richard Leakey of the National Museums of Kenya that "those who opposed the Southampton meeting will accept that they were wrong and a stronger world forum for archaeology will be the result".

Nigel Williams



Blood products

Hepatitis screening extended

Washington

SPURRED by growing concern that non-A and non-B hepatitis may represent a more serious health hazard than previously thought, the American Association of Blood Banks (AABB) announced last week that its members will begin screening all donated blood for evidence of non-A, non-B hepatitis. But as AABB officials are quick to acknowledge, such screening leaves much to be desired, as no direct tests for non-A, non-B hepatitis exists.

Blood banks hope to get an indirect indication of the potential for a donor to transmit non-A, non-B hepatitis by using two different blood tests. One measures the level of alanine amino transferase (ALT), commonly used as an indicator of liver dysfunction. The other tests for the presence of antibodies to hepatitis-B core antigen (anti-HBc). High levels of donor ALT and the presence of anti-HBc both correlate with subsequent development of non-A, non-B hepatitis.

The debate over whether to use one or both of these tests to screen donated blood has been raging for years. The ubiquity of non-A, non-B hepatitis as the cause of post-transfusion hepatitis became clear only after the development of a test for hepatitis-B surface antigen in the early 1970s. Between 7 and 17 per cent of all transfusion recipients develop post-transfusion hepatitis. Now, with effective screening for hepatitis-B, non-A, non-B hepatitis is estimated to account for more than 90 per cent of these cases. Even before such screening, Stephen Feinstone of the laboratory of infectious diseases at the National Institutes of Health (NIH) estimates that two-thirds of post-transfusion hepatitis cases were non-A, non-B.

The American Red Cross is also implementing ALT testing at its blood banks; the programme began on 7 July, and is expected to be completed by 1 October. AABB expects to implement testing in its member centres by 30 November. A third organization for blood banks, the Council for Community Blood Centers (CCBC) accounting for approximately one quarter of all blood collections, has not officially declared a position on ALT testing. But its president, Richard Counts of the Puget Sound Blood Center, says most members will go ahead with ALT screening.

Far more contentious is the use of anti-HBc testing. A study of post-transfusion non-A, non-B hepatitis in 481 recipients by Harvey Alter and associates at NIH showed (*Annals of Internal Medicine* 104, 488; 1986) that 11.9 per cent developed non-A, non-B hepatitis after receiving blood from donors positive for anti-HBc, compared with 4.2 per cent among recipients for blood negative for anti-HBc.

Alter estimated that prospective anti-HBc screening could eliminate 43 per cent of non-A, non-B hepatitis cases, while losing only 4 per cent of donors. Joseph Bove, chairman of AABB's committee on transfusion-transmitted diseases, argues that, taken with ALT screening, anti-HBc will go a long way towards reducing non-A, non-B hepatitis. Bove also argues that anti-HBc will be a useful "second line of defense" in screening for hepatitis-B. But Bove admits that there is no good biological reason to explain why people with antibodies to hepatitis-B core antigens should be more likely to transmit non-A, non-B hepatitis. The same positive relationship does not exist between hepatitis-B surface antibodies and transmission of non-A, non-B hepatitis.

Robert AuBuchon of the American Red Cross says the Red Cross is planning to start an anti-HBc screening programme of its own, but not until after the ALT test is implemented. Counts feels that the Food and Drug Administration should play a larger role in certifying the usefulness of anti-HBc. Counts also points out that Alter's study used a radioimmunoassay for anti-HBc testing, whereas an enzyme-linked assay is more likely to be utilized in a large screening programme. A further concern is that once implemented, screening tests are hard to disperse with. The Red Cross had planned a prospective clinical trial of non-A, non-B hepatitis screening techniques, but were daunted by the size of the project, and the possible ethical problems of giving un-screened blood to recipients.

What everyone is hoping for is a direct test for the agent causing non-A, non-B hepatitis, but that seems a long way off. Several candidates have been suggested, but none has held up. A major hurdle is tracking down the causative agent is its low concentration in blood — as much as 10,000 times less than hepatitis-B virus — and the lack of a suitable animal model for the disease. Only chimpanzees have been successfully infected with non-A, non-B hepatitis.

A major concern for all blood centres will be the loss of donors from false positives from the ALT and anti-HBc tests. Between 4 and 7 per cent of donors are expected to be prevented from giving blood. This comes at a time when screening for antibodies to the AIDS virus has already had a negative impact on the blood supply. The test will also increase the cost of a unit of blood by about \$3. Despite the additional cost and the loss of donors, AABB president Eugene Berkman says the tests are "essential to increase the safety of the blood supply".

Joseph Palca

Good prospects for lion tamarins

Washington

BRIGHTER days may lie ahead for the golden-headed lion tamarin (*Leontopithecus chrysomelas*). The National Zoo here in Washington recorded the first birth last

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

month in a captive breeding programme established to restore numbers of the endangered species.

Golden-headed lion tamarins have recently been the object of a worldwide conservation effort (see *Nature* 322, 586; 1986). The Washington infant's parents are themselves well travelled. They were part of a group of 16 golden-headed lion tamarins imported by a Belgian animal dealer in the winter of 1983 just before Belgium signed the Convention on International Trade in Endangered Species. The animals were ultimately repatriated to Brazil in November 1985 but were not released into the wild. Instead, an International Recovery and Management Committee for the tamarins, co-chaired by Jeremy J.C. Mallinson of the Jersey Wildlife Preservation Trust and Adelmair Coimbra-Filho of the Rio de Janeiro Primate Center, decided that the animals should be included in a captive breeding programme to be conducted both in Brazil and at several locations in North America.

Last March, a total of 20 golden-headed lion tamarins, including the 16 recovered from Belgium, were sent to the National Zoo in Washington and the Los Angeles Zoo. The National Zoo subsequently sent four animals to the Brookfield Zoo in Chicago. Under the terms of the breeding programme, all the US animals remain the property of Brazil, and may some day be included in a programme to reintroduce them to their natural habitat.

Devra Kleiman of the National Zoo, who is a member of the international management committee, says that until there is a better monitoring programme for wild golden-headed lion tamarins, it would be a mistake to release the captive population. In addition to the one birth, two other tamarins at the National Zoo appear to be pregnant. "Consider these as genes in the bank", says Kleiman.

Joseph Palca

Cameroon disaster

Carbon dioxide blamed

Washington

US INVESTIGATORS at Lake Nios, Cameroon, where 1,700 people were killed by toxic gases on 21 August, believe the disaster was similar to one that occurred in Lake Monoun, also in Cameroon, in 1984, which killed 37 people, according to the State Department Office of Disaster Assistance. There it is thought large quantities of carbon dioxide were released from the lake bottom. Both lakes are set in deep craters, and if the parallel proves correct there could be important implications for geochemically similar lakes in the area, including Lake Kivu.

Among the teams sent from the United States are limnologists, geochemists, pathologists and vulcanologists. Also in Cameroon is Harold Sigurdsson of the University of Rhode Island, who studied the Lake Monoun event. He and his colleagues concluded that carbon dioxide, possibly contaminated with other unidentified gases, was released following a disturbance of the lake bottom triggered by a landslide, perhaps caused by a minor earthquake. The carbon dioxide is thought to have slowly seeped from cold volcanic vents and accumulated over many years in high concentration, partly as bicarbonate, in the hypolimnion, along with ferrous ions reduced from siderite deposited in the lake by wind-blown loess. Nucleation points for the saturated solution were provided by the bottom disturbance and the gas came out of solution. Sigurdsson estimated that carbon dioxide was present in the hypolimnion at a pressure equivalent to 10 atmospheres; carbon-14 analysis showed 90 per cent of the carbon to be of volcanic origin.

Some observations of the bodies both at Monoun and Nios suggest that other gases may also have been present. Some descriptions speak of reddish skin burns and foaming at the nose and mouth that would not have been caused by carbon dioxide. Vegetation was also bleached, although clothing was unaffected. Sigurdsson and colleagues excluded sulphuric acid from oxidised hydrogen sulphide as a cause because it was virtually undetectable on the lake and because sulphuric acid usually causes black, rather than red, skin burns. Nitric acid may have been formed from oxidation of ammonia. The US teams will be taking samples from the water at different depths as well as from the atmosphere, and measuring temperature profiles. Pathologists may exhumate some bodies for examination; details will be presented to the US ambassador in Cameroon sometime this week or next.

Tim Beardsley

Museums

Science with a touch of magic

WHAT is Richard Gregory, head of the Brain and Perception Laboratory at the University of Bristol, doing bent over a snooker (pool) table? And how is it he never misses a shot?

Closer inspection will reveal that he has given himself an advantage — the inside of the table is elliptical. That means that if two balls are placed at the foci of the ellipses, the two will always collide whatever direction you hit one of them in.

The snooker table is one exhibit from Gregory's new museum, the Exploratory,

planes) up in the air.

Then there is a set of mirrors to lead you to a solution to the old paradox of why you look left/right reversed in a mirror but not upside-down. The answer is simple and difficult — according to Gregory, among those who have it wrong are Plato, Lucretius and Kant (and Martin Gardner in *The Ambidextrous Universe*.) The "exploratory" answer: that mirrors allow us to see the front of opaque objects although we are behind them: and the reversal is produced by the rotation of the object to face the mirror. You can try it with a book in front of a mirror: turn a page facing you to face the mirror either by turning it around or by turning it over. The latter produces no left/right reversal.

Dozens of exhibits force even trained scientists to think twice. Among them, Gregory himself is as excited as if he too is seeing them for the first time.

But how are children going to react? The full-size 20,000 square foot "Exploratory — Hands on Science Centre" will be receiving several school parties every day. Will they too recognize that "science has its own magic"? The Exploratory's director, James

Dalgety, who has spent the past fifteen years running a puzzle-manufacturing company, admits to few doubts.

Alun Anderson



which will open next month and from which a selection of exhibits are being previewed at this week's meeting of the British Association for the Advancement of Science in Bristol. Other than the Exploratorium in San Francisco, there is nothing quite like the new museum. It is a place for people to explore: it is not exhibits that do things when people press buttons, but people who do things with the exhibits until explanations emerge.

Many exhibits excite interest by producing effects that are counter-intuitive. For example, two tracks down which model cars can be run have been set up on one wall. One track is straight, the other curved. How many people will be willing to bet that the car on the curved track will reach the finish faster than that on the straight track? And how many believe that if you start a pair of cars off at different points on the curved track they will always arrive at the finish simultaneously wherever you start them from?

Elsewhere, a ball hovers six feet up in the air. It is suspended on a jet of air — no surprise until you see that the air jet is several feet away. Why does the ball not just drop out when the air stream is bent to one side? Some balls dangling on strings nearby and an ordinary hair-dryer give one a chance to explore the Bernoulli effect which keeps the ball (and aero-

- While Bristol's hands-on science exhibition concentrates on brain teasers, London's equivalent is more attuned to catching the imagination of the younger visitor. Under development for several years, "Launch Pad" was officially launched last week at the Science Museum. Most of the 70 exhibits are geared more to technology than to science. And the designers have usually chosen to provide a simple illustration of a principle rather than to set a problem.

Among the exhibits that may already have provided the formative experience for the next generation of technologists is the arch bridge: five blocks are arranged on supports to form the arch, the supports are removed and the builder is persuaded to walk over the bridge.

Where problem-solving is in evidence, it is likely to be in the form of assembling a lock and key or water pump from large parts in a transparent casing. The joke exhibit is a standard hot air hand-drier, an adjunct to the several popular exhibits that involve contact with water. Among these is a giant simulated toilet cistern and a miniature flow tank.

Peter Newmark

Nuclear power

India's reactors run into trouble

New Delhi

INDIA'S RAPP-1 nuclear-power reactor in Rajasthan state seems set to gain an unenviable world record as the first reactor to be prematurely retired. It has been in service for just eight years.

The Canadian-built 230 MW (E) pressurized heavy-water reactor has been out of action since March 1982 following detection of a crack in the end-shield of the reactor vessel. A four-year battle, using remote-controlled tools, failed to plug the leak in a highly radioactive and inaccessible area which, according to Dr M.R. Srinivasan, chairman of the nuclear power board, "has become embrittled by radiation". The only options now are an expensive operation to replace the end-shield after a long wait to "cool" the reactor, or a permanent close down of the £110 million reactor. According to nuclear engineers, closing down for good will be more economical.

More bothersome to India is the problem that has crippled the 100-MW (thermal) Dhruva reactor in Trombay, a key facility for producing plutonium. The high flux, £60 million reactor which went critical last August, was expected to generate about 60 kilograms of plutonium annually. But Dhruva has dashed India's hopes of producing and stockpiling plutonium because a design flaw prevents the reactor operating above 20 MW.

Dhruva, which uses natural uranium as fuel and heavy water as coolant and moderator, is a bigger and improved version of the Canadian-built 40-MW (thermal) reactor Cirus which supplied the plutonium for India's nuclear test in 1974. One major difference is that being a production reactor, Dhruva's fuel pellets are clad with aluminium (instead of the usual aluminium-zirconium alloy) to make reprocessing of plutonium easier. It is not known if the choice of aluminium for cladding and coolant tubes was a mistake. According to Dr P.K. Iyengar, director of the Bhabha Atomic Research Centre (BARC), Dhruva could not be operated at high power levels because of excessive vibration of the coolant tubes when heavy water was pumped through them under pressure. The vibrations could lead to rupture of the fuel rods inside the coolant tubes which would contaminate the coolant with fission products.

BARC has been trying to solve the problem for the past eight months and has now decided to dampen the vibrations by placing clamps at intervals along the fuel rods. According to Iyengar, simulated runs after partially loading the reactor with modified fuel bundles have shown satisfactory performance. But the question of whether the clamps will provide a

permanent solution will be answered only in actual operation. Similar vibration problems also cropped up in the Superphénix reactor in France according to Iyengar, who hopes that Dhruva can begin operation in a month.

Recommissioning of Dhruva is vital for India's goal of accumulating sufficient stocks of plutonium to fuel its breeder reactors in the second phase of its power programme. The 25-year-old Cirus, the only other plutonium-producing research reactor, is operating at half its capacity and is due for decommissioning. India has the means to reprocess plutonium from spent fuel, but international agreements

prevent it from doing so at four of its six power reactors. The United States has not permitted India to reprocess fuel from the two American-built reactors in Tarapur.

India has been reprocessing the fuel from the two 230-MW reactors in Rajasthan, RAPP-1 and RAPP-2, but an agreement with the Soviet Union (which supplied heavy water) forbids India from diverting the processed plutonium to any unrestricted facility. In a major move to overcome this restraint, BARC has now decided to burn the plutonium along with thorium in RAPP-2 itself. By conserving natural uranium, which is scarce in India, BARC says its strategy will help use the plutonium. India is free to use the plutonium processed from its power reactors in Madras but reprocessing of fuel has not yet begun.

K.S. Jayaraman

Environmental protection

Ice-minus test stopped yet again

Washington

THE first field test of bacteria that have been genetically engineered to protect plants against frost damage will not, after all, take place this year. The University of California, which is proposing the test, last month agreed with opponents that it would re-examine safety issues before going ahead. Because the test cannot be performed during the winter months, it is now unlikely to take place before next spring.

The test was first proposed by Steven Lindow of the University's Berkeley campus in 1982. Lindow plans to spray potato plants with a genetically altered strain of *Pseudomonas syringae* that has had removed the gene responsible for producing an ice-nucleating protein. Because the protein in naturally occurring *P. syringae* acts as a focus for ice crystal formation on host plants, Lindow hopes that the engineered form will reduce frost damage.

Lindow's experiment has already been the subject of environmental impact assessments and formal approval by two federal agencies, the National Institutes of Health and the Environmental Protection Agency. But earlier this month, local groups, together with Jeremy Rifkin's Foundation on Economic Trends, obtained a temporary restraining order against the university on the grounds that it may not have met the requirement of Californian law for a local environmental impact report if any significant impact is thought plausible.

The out-of-court agreement now reached obliges the university to review all the safety evidence once again and to decide within 30 days if a local impact report is indeed needed. But the plaintiffs have also agreed to provide any new evidence they want taken into account, and if they disagree with

the university's assessment have only a further 30 days in which to sue. Lindow says he hopes that the deal will mean that there will be no more last-minute hitches. Numerous and expensive preparations for the experiment had to be abandoned when the temporary restraining order was issued on 4 August, just 2 days before the test was due to start.

Tim Beardsley

Sun shines on UN's Kabul venture

UNITED Nations funds are helping to set up a new research and development centre for solar energy in the Afghan capital of Kabul, according to the official news agency Bakhtar. The foundation stone was laid on 5 August and construction is due to be completed by 1990.

Facilities are to include production workshops, laboratories, showrooms, small water stations, greenhouses, small animal husbandry farms, poultry rearing and a "typical biogas plant". Equipping the centre, and the training of "technical cadres" is expected to cost some US\$1.2 million. When completed the centre should meet the "energy needs of the people for lighting and heating of buildings, water supply and cooking".

This major UN investment represents a new phase in Afghanistan's development. Until the socialist "Sowr revolution", the principal source of foreign investment was the United States which, for example, had plans to develop at Kandahar the largest airport in Asia. Since the Soviet "fraternal intervention" of December 1979, the main source of aid has been the Soviet Union, which has provided the funds for facilities ranging from primary schools and hospitals to a sophisticated new broadcasting centre.

Vera Rich

RGO move FAR from OK

SIR—Your correspondents in the somewhat unfortunately titled letter "RGO move OK" (*Nature* 322, 402; 1986), claim that *Nature* does not fully realize the implications of the move of the Royal Greenwich Observatory (RGO) to Cambridge. This shortcoming is clearly not unique.

They state that the move to Cambridge will consolidate research astronomy with instrument science and engineering. Yet this ignores both the stated policy of SERC and its recent history in directing the affairs of RGO.

When many people now working in high technology at RGO were recruited, they joined a national observatory in its own right. The research programme was then considered an important part of the life of the observatory, while traditional long-term projects were considered of value and were pursued vigorously. Although instrumentation at that stage had not kept up with the other work, a very positive environment allowed and encouraged the introduction of the advanced instrumentation and engineering techniques essential to an international class observatory.

Over the past few years, however, SERC has sought to erode this by a process of attrition: it has restricted the astronomical research role to ten per cent of manpower effort; it has continually cast doubt on the long-term projects being carried out here; and now, in its current forward look, it has established firm plans to cut severely the applied science and engineering aspects of the work. The result will be that manpower allocated to this area in 1990 will be around one-third of its 1982 level, of which a significant fraction will be allocated to La Palma for maintenance support. Clearly the capability of RGO in 1990, particularly in instrumentation and technology, will be nothing like as great as it is today.

It has been stated by SERC that the move must be self-financing. This can be achieved only by selling the Herstmonceux site, which is immeasurably more valuable to SERC than its sale price can possibly be. Even then, the move cannot be made self-financing without cutting manpower dramatically or greatly reducing the facilities available. It appears that SERC plans to follow both courses of action and will impose a burden on the university to make up for the shortfall in support.

This curious view of economics by SERC is not confined to RGO. The recent withdrawal from the South African Astronomical Observatory and near-withdrawal from the Anglo-Australian Observatory (AAO) have similar implications. Had the AAO withdrawal occurred (and the agreement must still be consi-

dered under threat), SERC would have spent about £30 million on La Palma without any significant corresponding increase in the observing time available to UK astronomers. The generosity of SERC in providing capital facilities, which you so rightly mention in your leading article of 3 July (322, 1; 1986), could surely have been put to better effect when considering the resources necessary to run those facilities that already exist.

What makes the prospect of a move to Cambridge so dismal is not just the upheaval and waste of effort involved, but the fact that there will be no change in style of management by SERC. There will be the same unwillingness to state aims and to provide the resources to execute those aims (or to reduce them correspondingly); there will be the same cavalier attitude to staff morale; there will be the same excessive concentration on capital facilities at the expense of running existing facilities to produce science.

It therefore seems likely that what the University of Cambridge expects and what it receives will not correspond.

We strongly urge the chairman and council of SERC to reconsider their decision in order to minimize the serious damage already done. Perhaps we could then sit down together, in the democratic manner that you mention in your leading article, and discuss long-term aims that would offer British astronomy the best prospects for the future; council's plans clearly do not do that at present.

I.G. VAN BREDA & P.D. READ

Royal Greenwich Observatory,
Herstmonceux Castle,
Hailsham, East Sussex BN27 1RP, UK

SIR—Contrary to your headline-writer's caption to the letter (*Nature* 322, 402; 1986) about the proposed move of the Royal Greenwich Observatory (RGO) to Cambridge, the RGO move is NOT OK. The move to Cambridge is merely the least of several evils — but the Science and Engineering Research Council (SERC) may also have miscalculated the cost.

The rumour mill suggests that the cost of moving RGO is roughly £6 million and SERC claims that the move can be self-financing. However, cursory research will uncover various press articles on the current slump in the price of castles (see, for example, an article in the *Sunday Times* of 13 July). Apparently "Fort Belan, an 18th century fort complete with cannon, drawbridge and ramparts guarding the Menai Straights in North Wales has still not found any takers (although it has been offered) complete with an airfield, three miles of coastline and a dockyard for £750,000."

Agricultural land, the property associated with Herstmonceux Castle, is at its lowest value for years, partly due to changes to the agricultural support policies of the European Communities and partly amplified by the dumping of land by city speculators. I conclude that it is highly improbable that the sale of the castle will fetch much more than £0.5 million. I fear that SERC will be committed to the move before its illusions about the value of the Herstmonceux site are shattered. Having to find £5 million will pauperize UK astronomy and further reduce SERC's funding of other sciences.

Either SERC must experimentally establish the value of Herstmonceux Castle by putting it on the market before it takes irrevocable steps or the Treasury must reject this "self-financing" proposal as based on an untested theory. Factor-of-ten errors may be acceptable in astronomy but not in the financial provision for it.

MICHAEL PENSTON

Rose Lodge,
Magham Down,
Hailsham BN27 1PR, UK

Fish-eating Eskimos?

SIR—Your Washington correspondent Tim Beardsley's account of the US nutritional research scene¹ refers to the "low rate of heart disease among Eskimos who eat a lot of fish". But of the sixteen major Eskimo groupings in the western Arctic zone², only three (Aleut Eskimos, Southern Alaskan and Western Alaskan Eskimos) ingest fish as a portion of their daily diet. Most consume lipids of marine mammalian origin.

The beneficial effects of omega-3 polyunsaturated fatty acids derived from marine sources were reported by Dyerberg's group³⁻⁵ based on their studies on Greenland Eskimos who predominantly consume whales and seals, but not fish. Bang *et al.*³ reported that the frequency score of meals (average number during one week's consumption) eaten by Greenland Eskimos decreased in the following order; seal meat and blubber 6.4; whale meat and blubber 5.7; soup with seal meat 2.3; fish 1.4; and seal intestines 0.6. It is also interesting to note that the word Eskimo is apparently derived from a Cree Indian word, meaning "eaters of raw meat".

SACHI. SRI KANTHA

Laboratory of Marine Biochemistry,
Faculty of Agriculture,
University of Tokyo,
Bunkyo-ku, Tokyo 113,
Japan

1. Beardsley, T. *Nature* 322, 104 (1986).
2. Carter, W.K.C. *Encyclopedia Britannica* Vol. 1, 15th edn. 1126 (Encyclopedia Britannica Inc., Chicago, 1983).
3. Bang, H.O., Dyerberg, J. & Hjorne, N. *Acta med. Scand.* 200, 69-73 (1976).
4. Bang, H.O., Dyerberg, J. & Sinclair, H.M. *Amer. J. Clin. Nutr.* 33, 2657-2661 (1980).
5. Dyerberg, J. *Nutr. Rev.* 44, 125-134 (1986).

Have the heavy neutrinos gone?

The pursuit of particle physics without the use of accelerating machines is always an intriguing challenge. But the simplicity of measurements may be offset by complications of interpretation.

THE search for the heavy neutrino continues, but the chances of success seem steadily to diminish. That is the curious psychological effect of a demonstration, just published, of the pitfalls ahead of those who look for evidence of heavy neutrinos (where "heavy" merely means "greater than zero") in the electron spectrum associated with beta-decay.

The argument, due to J. Lindhard and P. G. Hansen from the University of Aarhus in Denmark (*Phys. Rev. Lett.* **57**, 965; 1986) is interesting not so much because it is decisive, but because of its simplicity and clarity. Its bearing on the question whether neutrinos can have even a modicum of mass is only indirect.

The idea that neutrinos may have mass is now some 15 years old. The successful unification of the weak and the electromagnetic force-fields implies that there should be as many kinds of neutrinos as there are different kinds of electron-like particles, which for the time being means three (the tauon and muon with the ordinary electron). The notion that some neutrinos may have mass arises from attempts to unify strong and weak nuclear forces.

Cosmologists do not bother to conceal their delight at the prospect that a nearly universally distributed particle might carry enough mass to help close the Universe gravitationally. Everybody seems to agree that, if there are indeed three kinds of neutrinos with very similar properties, those actually observed will be mixtures (in the quantum sense of being linear combinations of fundamental states) of the three basic neutrinos. Specifically the neutrinos (strictly, anti-neutrinos) emitted in partnership with electrons during beta-decay will on the average carry a fraction of any mass there may be.

This was the inspiration of an accurate measurement 10 years ago by Soviet physicists of the energy-spectrum of electrons emitted in the beta-decay of tritium. The point is merely that if the associated neutrinos have even the smallest amount of mass, the requirement that energy and momentum should be conserved in the decay will distort the energy spectrum that would otherwise be expected.

As neutrinos as such cannot be observed directly because they hardly interact with matter, the predicted distortion of the electron energy-spectrum may offer the best hope, for the time being, of measuring their properties. A quick calculation will show that there will be a ten-

dency for an excess of electrons at low energy.

Tritium is an obvious place to start, if only because the transition energy is relatively small and because the effect on the decay process of the atomic electron shell can be calculated easily, given the simplicity of the structure. But the same search for distorted energy spectra has been carried out with the isotopes ^{64}Cu , ^{35}S and no doubt others.

Most experiments have yielded only inconclusive evidence. Datar, V.M. *et al.* (*Nature* **318**, 547; 1985), for example, working with ^{35}S , found that the admixture of a single supposedly massive neutrino corresponding to, say, a muon with the supposedly zero-mass electron neutrino would be, at most, a mere 1 per cent.

The outstanding result in the field remains that of J.J. Simpson, from the University of Guelph in Ontario, who has studied tritium decay by a neat technique in which tritium atoms are embedded in a silicon detector that functions as a means of measuring the total loss of energy by ionization within the silicon, for which purpose the semiconductor is doped with lithium (*Phys. Rev. Lett.* **54**, 1891; 1985).

Simpson's conclusion, going against the grain of earlier evidence except that from the Soviet Union, is that the energy spectrum of the decay electrons does indeed differ from that expected on the simplest principles, and in such a way as to suggest that the massless neutrino is mixed to the extent of 3 per cent with a heavier neutrino whose mass is the equivalent of approximately 17.2 keV. Ungratefully, even cosmologists are not enthusiastic: this neutrino is embarrassingly massive.

Lindhard and Hansen make one elementary point about Simpson's experiment, that it differs from most others of this kind in that the decaying tritium nuclei embedded in silicon are mostly stripped bare of their single electron. It follows that the decay energy is not the conventional 18.6 keV, the difference of energy between an atom of tritium (with one electron) and an atom of its decay product, ^3He , which has two electrons.

Most simply, what this remark implies is that the energy available in the decay of a stripped tritium nucleus is less than that which might be wrung from the decay of an intact atom by the difference between the absolute values of the two ground-state electron energies, both negative (of which that of helium is larger to the tune of

more than 65 eV). But to say that the energy available will be different for the decay of a stripped and for an intact atom does not imply that all the difference will be conferred on the decay electrons, although the magnitude of the effect is such as to suggest that Simpson's small differences could easily be swamped.

What Lindhard and Hansen have done is to make qualitative calculations of how the difference of available energy between stripped and intact atoms will be reflected in the energy spectrum of the electrons. The most obvious influence is that of the atomic electron in an intact tritium atom, which will have the effect of increasing the outward energy of the electron. Other effects of the atom on the departing electron are more subtle.

The fact that the sudden conversion, in beta-decay, of a tritium nucleus to a helium nucleus leaves the lone electron in an excited state implies that some part of the 18.6 keV difference of energy between the two ground states is released not to the electron or neutrino, but in de-excitation. Similarly with the deionization that must follow that adjustment. Note that the term "sudden" is not merely descriptive but is a technical term in quantum mechanics.

The calculation of the quantities involved is unavoidably messy; that for the screening correction was first done by M.E. Rose in the 1930s. Lindhard and Hansen add a nice twist by using a function to represent the screening potential of an electron shell that includes an exponential term yet nevertheless allows for exactly soluble wave functions.

The upshot is the conclusion that the energy carried off by electrons will be decreased, on the average, by some 34 eV, or that the expected electron energy spectrum calculated from the full value of 18.6 keV will be shifted bodily towards the origin of energy by that amount. This, the authors claim, is enough to account for two-thirds of the discrepancy claimed by Simpson. Chemical interactions, weighing in on the same side as de-excitation and deionization, could account for the rest.

There is no logical reason why this should strengthen beliefs that massive neutrinos do not exist. The theories are explicit on the point, while there are many other types of experiments in which the issue may be tested. But there is a sense in which negative trials of a hypothesis count positively against it: many people are prejudiced against failure. **John Maddox**

Cell biology

Is there a common design for cell membrane channels?

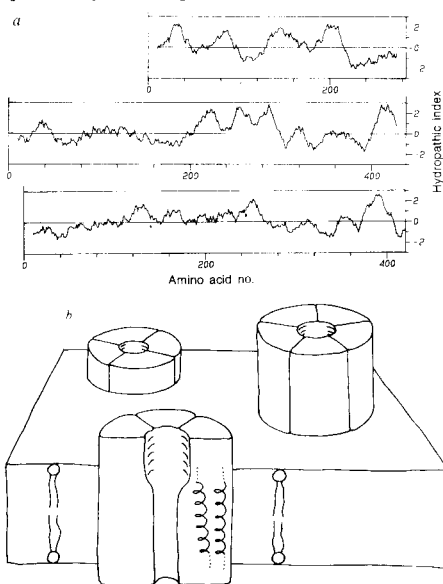
from Nigel Unwin

THE first membrane channels to be seen clearly in electron-microscope images were those at gap junctions, described by David Robertson¹ a little more than 20 years ago. The molecular characterization of these channels has recently taken an important step forward, as the primary sequence for a gap-junction polypeptide, reported in two of the latest issues of the *Journal of Cell Biology*^{2,3}, has now been deduced from complementary DNA clones. This accomplishment follows closely the analyses in Shosaku Numa's laboratory of the voltage-dependent sodium channel⁴ and the subunits of the nicotinic acetylcholine receptor⁵, making a total of three eukaryotic cell membrane channels whose sequences have been derived by recombinant DNA methods. Intriguingly, all three channels are major participants in the transmission of electrical signals in excitable tissue. Given that they are functionally related, although individually quite distinct, is it possible that they are all simple variations of a common structural theme? Several hints, or analogies with other systems, suggest they could be.

First, the new details about gap junctions. In David Paul's experiments² the cDNA encoding a gap-junction polypeptide was obtained by screening a rat liver library with an antibody specific to the gap junction, whereas Nalin Kumar and Bernie Gilula³ used a human liver library and a synthetic oligonucleotide probe based on the previously determined amino-terminal sequence⁶. In either case, the derived primary sequence corresponds to essentially the same polypeptide (only 4 residues difference between human and rat), which has a relative molecular mass of 32,000 (32K). This polypeptide contains no amino-terminal signal sequence and may correspond exactly to the channel subunit normally detected on SDS gels as a 28K band, although there could, for example, be post-translational cleavage of a carboxy-terminal domain. The three-dimensional structure of this polypeptide, when the time comes, will be doubly interesting: not only must it assemble an ion channel in the membrane of one cell, but also it needs to link tightly to a channel in the neighbouring cell so that the pair together create a continuous, leak-proof communication pathway between the cell interiors.

To those of us who have scrutinized hydropathy profiles for the subunits (or domains — see below) of the other two

channels, the profile for the gap-junction polypeptide (*a* in the figure) looks remarkably déjà-vu. In common with the other channels, there are several strings of hydrophobic amino acids around 20 residues long. Such sequences are uncharacteristic of soluble proteins, but probably correspond to transmembrane



By accident or design? *a*, Similar hydropathy profiles¹² (19-amino-acid window) for the rat liver gap-junction polypeptide (top); α -subunit of the acetylcholine receptor (middle); and amino-terminal portion of sodium channel containing the first homologous domain (bottom). *b*, Diagram emphasizing the apparent structural similarities of the three channels.

α -helices, as direct sequence and structural comparisons of the L and M subunits of the photosynthetic reaction centre convincingly demonstrate⁷. Alpha-helical models for the acetylcholine receptor and sodium channel have been derived on this assumption and clearly the implication now is that the transmembrane portion of the gap-junction polypeptide also contains several α -helices. Thus I think that the original ' α -helical' interpretation of the X-ray scattering pattern from isolated gap junctions⁸ is closer to the truth than the ' β -sheet' interpretation quoted in ref. 2. Indeed, it looks as if the α -helical bundle is a structural motif common to the bilayer portion of all the channels, whatever the specialized features required to meet their specific physiological needs.

A fundamental aspect of molecular design is the symmetry. For each of the channels the symmetry has either been evaluated experimentally or (in the case of

the sodium channel⁴) deduced from the sequence. The molecules each consist of like units arranged in the plane of the membrane around a central aqueous pathway, the pore. The gap-junction channel is made from a ring of six identical subunits; the acetylcholine receptor from a ring of five similar (including two identical) subunits; and the sodium channel from a ring of four similar domains composing about 60 per cent of a single polypeptide chain. Evidently the gap-junction channel is the only channel with true rotational, or cyclic, symmetry; however, the amino-acid homology between the similar subunits (about 40 per cent) or between the similar domains (about 50 per cent) is certainly high enough to imply that the corresponding tertiary and secondary configurations are largely the same. Hence the respective structures would be expected to have good symmetry at this level over much of their length, as is indeed suggested by direct images of the acetylcholine receptor⁹.

Thus, all three channels apparently possess cyclic symmetry, or a very good approximation to it, the symmetry axis delineating the pore through the membrane (*b* in the figure). Is this common architectural plan reflected in other features? Considering the parallels noted in the internal structure of the proteins, it may be no coincidence that the constricting diameters of the pores — about 16 Å for the hexamer (gap junction), about 7 Å for the pentamer (acetylcholine receptor) and about 4 Å for the tetramer (sodium channel) — vary in proportion, qualitatively, with the number of elements around them. It is also interesting that the pores of the gap-junction channel and acetylcholine receptor seen in electron-image maps^{9,10} are of the same character, constricted at one end (cytoplasmic) but wide open for an extended length at the other. As for the implications, perhaps the similar structural patterns are clues that the channels respond to chemical or electrical stimuli by related mechanisms. Except for the energetic constraint that subunit displacements should be predominantly parallel, rather than perpendicular, to the plane of the bilayer, there seems to be no reason why the rules should not be much the same as those for a soluble multi-subunit complex, such as haemoglobin.

Simple variations of a common structural theme occur frequently in nature. The protein shells of spherical viruses as diverse as those which infect plants or give us colds are, it turns out, such an example (see the recent discussion in *News and Views*¹¹), and may represent the type of assembly to which the channels are most analogous. Tomato bushy stunt virus, for instance, like the gap-junction channel, is constructed from multiple copies of the same polypeptide; human cold virus, like

the acetylcholine receptor, is constructed from different but related major polypeptides; satellite tobacco necrosis virus is constructed from one species of polypeptide, but is a smaller virus, having a third as many copies as the others. The striking discovery to emerge from a series of outstanding X-ray crystallographic studies conducted over the past few years is that the structures of the major polypeptides of all these viruses are very similar. Animal or plant virus, big or small, the basic motifs are much the same.

Urgently needed now from at least one of these channels is high-resolution information of the sort obtained from the viruses. Substantial technological advances have been made in macromolecular crystallography and membrane protein crystallization over the past few years, so it should not be too long before that dream is realized. The details revealed would provide a three-dimensional framework for rational site-directed mutagenesis experiments and other studies aimed at

probing precise physiological functions. They should also yield important insight into the general structural principles involved in switching between open, closed and inactivated (or desensitized) states. Contrary to the impression given by the variety of current models, it could be that all membrane channels work in much the same way. So far, at the junction between structure and function, there is a bit of a gap! □

1. Robertson, J.D. *J. Cell Biol.* **19**, 201 (1963).
2. Paul, D.L. *J. Cell Biol.* **103**, 123 (1986).
3. Kumar, N.M. & Gilula, N.B. *J. Cell Biol.* **103**, 767 (1986).
4. Noda, M. *et al. Nature* **312**, 121 (1984).
5. Noda, M. *et al. Nature* **302**, 528 (1983).
6. Nicholson, B.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 7594 (1981).
7. Michel, H. *et al. EMBO J.* **5**, 1149 (1986).
8. Makowski, L., Caspar, D.L.D., Phillips, W.C. & Goodenough, D.A. *J. Cell Biol.* **74**, 629 (1977).
9. Brisson, A. & Unwin, P.N.T. *Nature* **315**, 474 (1985).
10. Unwin, P.N.T. & Zampighi, G. *Nature* **283**, 545 (1980).
11. Wiley, D.C. *Nature News and Views* **317**, 108 (1985).
12. Kyte, J. & Doolittle, R.F. *J. molec. Biol.* **157**, 105 (1982).

Nigel Unwin is in the Department of Cell Biology, Stanford University School of Medicine, Stanford, California 94305, USA.

Optical computing

Elements of an optical engine

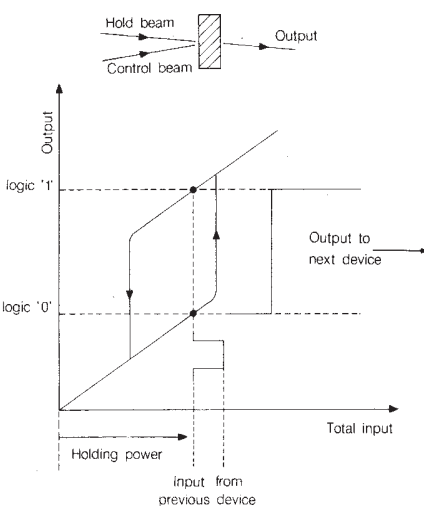
from Alan Miller

CAN the massive parallelism across an image and speed of light be harnessed to create a computer to compete with an already highly developed electronics technology? Two recent announcements bring an answer considerably closer. Heriot-Watt University in the United Kingdom and AT&T Bell Laboratories in the United States have now both demonstrated early cascaded bistable arrays — an essential requirement for digital optical parallel computing.

Digital optics is at an early stage of development, but the goal of both groups is to develop the technology required to demonstrate that optics can make a significant impact in digital processing and computing, rivalling silicon or gallium arsenide circuits in specific applications when and where electronics becomes limited. The advantage of optics is that whereas electrons interact strongly with each other, photons pass through each other unperturbed in the absence of any nonlinear interaction. This makes light a powerful, low cross-talk, high bandwidth means of interconnecting devices in parallel by using holograms, fibres or simply novel geometrical optics, which is beginning to be exploited for computers. But to go beyond mere communication and create an all-optical processor, optical logic and memory have had to be developed.

The first semiconductor optical equivalents of the transistor were demonstrated seven years ago in simultaneous research

at Heriot-Watt University¹ (using InSb) and at Bell laboratories² (using GaAs) that exploited newly discovered band-gap resonant, giant optical nonlinearities in semi-conductors³. By applying optical feedback in the form of reflecting mirrors



An optical logic element depends on a 'hysteresis' between input and output light. With the element in its opaque (logic '0') state, a hold beam keeps the device close to a transition to transparency (logic '1'). A small input signal can then flip the device through the transition and let the hold beam through. For cascaded arrays the power thus transmitted must be intense enough to act as a switching beam in subsequent elements.

on either side of the nonlinear crystal (often just the reflecting surfaces of the semiconductor), all-optical bistable or memory devices (see figure) were achieved that need only a few milliwatts of optical holding power⁴. All-binary logic operations have since been demonstrated. Optical bistability in other semiconductors has been actively pursued in various laboratories in the United States and Europe.

Successful implementation of these devices in digital optics will impose stringent requirements on their operation, including stability and reliability; low-power operation at relatively high speeds; good contrast between low and high states; and the ability to pack arrays of devices densely. Any one logic element must be able to accept inputs from several other devices (fan-in) and be able to communicate its output to several more (fan-out). The cascability of elements depends on the gain of the device, that is, the ratio of the change of throughput of the holding beam, to the additional input above the holding power required to switch (the holding beam functions as the power supply).

The group at Heriot-Watt had previously demonstrated cascability by sequentially coupling the output of one InSb device to the input of a second device on the same crystal such that the combination gave an XNOR gate⁵. Its recently announced work, which goes as far as producing the first cycling logical machine (the guts of a computing engine) forms part of the Joint Optical Bistability project (EJOB) supported by the Commission of the European Communities. EJOB links 12 laboratories with the aim of demonstrating a primitive optical computer. Under this scheme the Heriot-Watt group has now demonstrated an optical circuit consisting of three separate ZnSe bistable plates positioned in a loop such that green beams from an argon-ion laser can be passed from one plate to the next around the loop. Holographic elements divide the beams into arrays of three spots on each plate while signal and clocking pulses are provided by acousto-optic modulators. This shows the transfer of data, in parallel, around a processing loop in a 'lock and clock' sequence, an inversion occurring on each full cycle of the loop — essentially the primitive cycling computer that is the aim of EJOB.

The bistable plates, the essential nonlinear optical elements used for this demonstration, are thermally evaporated, multilayer band-pass interference filters containing ZnSe as the nonlinear spacer layer⁶. Optical bistability at a few milliwatts is achieved by thermally induced refractive index changes. These devices have the appeal of being relatively easily constructed with a high degree of uniformity over large areas, although some

stability problems need to be overcome.

In contrast, the Bell laboratories work, announced at a conference on lasers and electro-optics in San Francisco this June, requires very high-technology, ultrathin multilayers of GaAs and GaAlAs, each only a few atomic layers thick, grown by molecular beam epitaxy to a total thickness of a few micrometres⁶. The Bell group used an electro-absorption effect so the feedback necessary to achieve optical bistability is electrical instead of all-optical; the arrangement is therefore hybrid. The novelty is that the electrical feedback is intrinsic to the device, as the material acts as both photoconductor and optical modulator, so only a constant electrical bias need be applied. Furthermore, a built-in bias resistor is controlled by an auxilliary light source.

By making use of the anomalously large electric field shift of a quantized exciton in GaAs layers less than 100 Å thick (quantum confined Stark effect⁶), the Bell group showed optical bistability could be achieved at between 1 mW and 1 pW in the 834–859 nm range in these self-electro-optic effect devices (SEEDs). The relatively low switching powers are traded, however, for switching speeds in the range 1 µs to 10 s by altering the bias conditions with a He–Ne laser. A 2×2 array was made by etching 200 µm square mesas. Improvements to the performance of these devices allowed more than a 2.5:1 contrast ratio with a high degree of uniformity over the four devices, thus offering cascaded bistability in devices compatible with existing GaAs optoelectronics technology.

Both these examples of cascable arrays use devices that need much more refinement in terms of number of elements, speed and energy consumption before they could be practical, but it is worth bearing in mind that even a 100 × 100 array of devices switching at 1 µs would offer an impressive 10¹⁰ gate switches per second. What these demonstrations do prove is that semiconductor components for photonic switching and digital parallel processing now exist. Both can be readily scaled to accommodate larger numbers of elements offering test beds for exploring innovative ideas on architectures most suitable for parallel

arrays of gates.

The first applications of these optical switches may well come from routing in fibre communications where the input from many thousands of fibre channels is already in a highly parallel form. Optical switching arrays in a pipelined sequence could route data using coded pulses, thus avoiding the intermediate conversion to electronics which will become increasingly more difficult as the operating frequencies

increase¹⁰. The algorithms used in this application are not too far removed from many computing operations. Other applications appropriate to parallel processing would be, for instance, image processing, machine vision and radar array processing, which are already stretching electronic systems to their limits. □

Alan Miller is at the Royal Signals and Radar Establishment, Malvern WR14 3PS, UK.

Plant ecology

Counting the costs of rainfall

from Peter D. Moore

PRECIPITATION brings both benefits and problems to the plant life of an ecosystem, as demonstrated by several recent reports. Rainfall provides a source of nutrients which may be critically important to the recovery of some ecosystems when they experience catastrophes, or, as in the pine savannas of Central America, human clearance¹. But rainfall can also damage the plant canopy, either by physical impact² or by leaching elements from the leaves³, both of these processes resulting in decreased productivity, especially if the rainfall is acidic⁴.

There are many well-known examples of precipitation acting as a major source of nutrient capital for a growing ecosystem. Art *et al.*⁵ documented the growth of scrub woodland on a sandy barrier island which is dependent on and limited by the input of nutrients from the oceanic spray and precipitation of the northeastern coast of the United States. Similarly, Chapman⁶ showed that rainfall in Britain provided enough nutrient material (with the possible exception of phosphorus) in a 12-year period to account for the recovery of heathland after regular fire. In both these sites the substrate is sandy and nutrient poor, so that soil reserves are generally low and weathering is not a rich source of new material. Under these circumstances the precipitation input is especially important for the replenishment or growth of the biomass nutrient reserve.

Similar problems are found in the *Pinus caribaea* pine savannas of Central America, recently described by Kellman and Carty¹. These areas were once subjected to fires, but are now managed for forestry, which places the ecosystem under a further nutrient stress as timber is collected from it. Kellman and Carty estimate that the first plantation of pines creates a greater demand for plant macronutrients than can be supplied by the rainfall alone. Considering the first 30 years of growth, they show that sufficient potassium is supplied by rain (3.4 kilograms per hectare per year) to account for the biomass growth, but for calcium only about

60 per cent, for magnesium 30 per cent and for phosphorus 40 per cent of the requirements of the forest are met. Because these data refer to above-ground biomass only, this is undoubtedly an overestimate of the proportions supplied. Clearly the trees are drawing on the soil reserves during the first 30 years of growth. But in subsequent years, as harvests are taken, a steady state of nutrients can be maintained if forest management permits only the extraction of boles while the branches and roots are allowed to rot on site. The nutrient return from this source, coupled with the rainfall input, should account for further forestry needs.

But the passage of rain through a forest canopy can have negative as well as positive effects. Early work by Carlisle, Brown and White⁷ showed that in Britain considerable quantities of nutrients are leached from the canopy by the passage of rainfall, especially in spring and in autumn. The bulk lost in this way is often of a similar order to that contained in the deciduous leaf litter fall. An important further question is whether canopy leaching also results in substantial losses of organic matter, part of the primary productivity of the system. The data of Carlisle *et al.* suggested that it did, but the authors could not separate the effects of aphid secretions of honeydew in their water collections from those of direct leaching.

The new study of Amthor³ directly addresses this problem by analysing data from the Hubbard Brook forest in the United States and from two Himalayan forests. Considering only the summer growing season of the temperate forest and the monsoon period of the Indian forests, Amthor concludes that carbon losses amount to only about 1 per cent of net primary productivity in the temperate site and rather less than that in the tropics. So rain-leaching losses of carbon compounds do not represent a significant proportion of primary productivity.

Yet the detrimental influence of rain, particularly acidic rain, on the overall yield of many plants is now well estab-

1. Miller, D.A.B., Smith, S.D. & Johnston, A. *Appl. Phys. Lett.* **35**, 658 (1979).
2. Gibbs, H.M. *et al. Appl. Phys. Lett.* **35**, 451 (1979).
3. Miller, A., Miller, D.A.B. & Smith, S.D. *Adv. Phys.* **30**, 697 (1981).
4. Gibbs, H.M. *Optical Bistability: Controlling Light with Light* (Academic, Orlando, 1985).
5. Smith, S.D. *Nature* **316**, 319 (1985).
6. Smith, S.D., Mathew, J.G.H., Yaghizadeh, M.R., Walker, A.C. & Wherrett, B.S. *Opt. Commun.* **51**, 357 (1984).
7. Miller, D.A.B., Henry, J.E., Gossard, A.C. & English, J.H. in *Conference digest, Conference on Lasers and Electro-Optics* (San Francisco, June 1986).
8. Miller, D.A.B. *et al. IEEE J. quant. Electron.* **QE-21**, 1462 (1985).
9. Huang, A. *Phil. Trans. R. Soc. A* **313**, 205 (1984).
10. Midwinter, J.E. *IEEE Proc.* **132 J** 371 (1985).

lished. Recent data on soybeans⁴, for example, using simulated acid rain, has shown that a reduction in pH from 5.6 to 4.4 can cause an 11 per cent reduction in seed yield for the species. An influential process in the depression of growth may be the destruction of the wax cuticle of leaves by the direct impact of rain droplets, as demonstrated by Baker and Hunt². These workers observe the effects of experimental water droplets travelling at low to medium velocity (0.25–5 metres per second) using scanning electron microscopy and record the stripping of epicuticular wax from the leaves of *Brassica*, *Pisum* and *Eucalyptus*. Such removal of the protective, hydrophobic coat of leaves

could clearly have unpleasant side effects in leaving the plant tissues exposed to pollutant chemicals, pesticides and pathogens. This is the unacceptable face of the soft, refreshing rain. □

1. Kellman, M. & Carty, A. *J. appl. Ecol.* **23**, 211 (1986).
2. Baker, E.A. & Hunt, G.M. *New Phytol.* **102**, 161 (1986).
3. Amthor, J.S. *New Phytol.* **102**, 359 (1986).
4. Evans, L.S., Lewin, K.F., Owen, E.M. & Santucci, K.A. *New Phytol.* **102**, 409 (1986).
5. Art, H.W., Bormann, F.H., Voigt, G.K. & Woodwell, G.M. *Science* **184**, 60 (1974).
6. Chapman, S.B. *J. Ecol.* **55**, 677 (1967).
7. Carlisle, A., Brown, A.H.F. & White, E.J. *J. Ecol.* **54**, 87 (1966).

Peter D. Moore is a Reader in Ecology in the Department of Biology, King's College London, 68 Half Moon Lane, London SE24 9JF, UK.

Immunology

The growing immunoglobulin gene superfamily

from Tim Hunkapiller and Leroy Hood

MEMBERSHIP of the immunoglobulin gene superfamily, named for the immunoglobulin light- and heavy-chain gene families, has been growing rapidly in recent years¹. The report by Johnson and Williams on page 74 of this issue² on the gene encoding a rat T-cell-associated antigen adds yet another member to the expanding list. Members of this superfamily all share a common structure called the immunoglobulin homology unit³, a structure composed of a sequence about 100 amino-acid residues long and characterized by a centrally placed disulphide bridge that stabilizes a series of antiparallel β strands into the so-called antibody fold⁴. The variable (V) and constant (C) homology units, defined for the respective portions of the immunoglobulin chains from which they were identified, have similar but distinct three-dimensional structures. Members of the superfamily extend beyond the immune system and it appears the homology unit has played a central role in the evolution of cell-cell recognition.

The immunoglobulin gene superfamily so far includes eight multigene families and twelve single-gene representatives (see figure). The multigene families include the light (λ , κ) and heavy-chain gene families of immunoglobulin, the α , β and γ families encoding T-cell receptors and the class I and class II genes of the major histocompatibility complex (MHC)¹. Single-gene members include those encoding T-cell accessory molecules involved in class I (CD8)^{5–7} and class II (CD4)^{8,9} MHC recognition and possibly ion channel formation (T3 δ , T3 ϵ)^{10,11}; a receptor responsible for transporting certain classes of immunoglobulin across mucosal membranes (poly-Ig)¹²; β_2 -

microglobulin, which associates with class I molecules¹³; a human plasma protein with unknown function (α IB-glycoprotein)¹⁴; two molecules of unknown function with a tissue distribution that includes both lymphocytes and neurones (Thy-1, OX-2)^{15,16}; and two brain-specific molecules, N-CAM and neurocytoplasmic protein 3 (NP3)^{17,18}. These widely divergent examples indicate the incredible evolutionary versatility of the immunoglobulin homology unit.

CD8 is a homodimer in humans and a heterodimer in rats and mice (containing the Lyt 2 and Lyt 3 chains). The primary structure of the human CD8 single chain

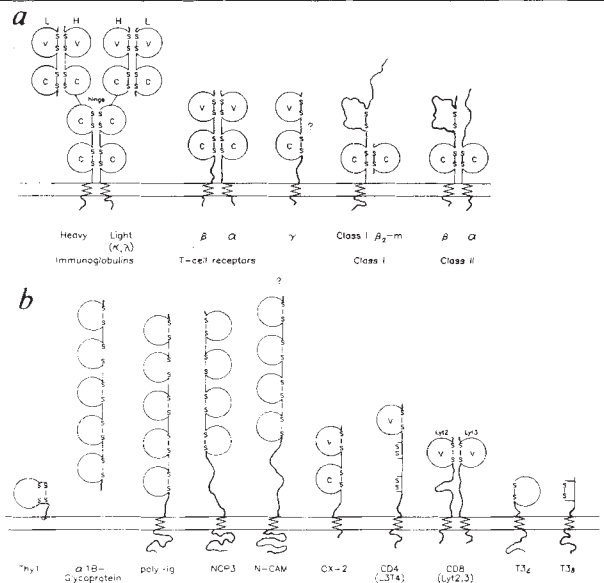
(also denoted T8 or Leu2) and its homologue in mouse (Lyt 2) and rat (OX-8) have been previously determined and found to contain one V homology unit, most like those of light chains, as well as a connecting hinge sequence and transmembrane and cytoplasmic domains^{5–7}.

In this issue, Johnson and Williams² complete the molecular characterization of rat CD8 with the report of the complementary (c)DNA sequence of the Lyt 3-like chain. This chain also has one V-like homology unit, consistent with the notion that CD8 recognizes conserved determinants of class I molecules with a receptor domain similar to the paired homology unit, antigen-binding domains of immunoglobulin and the T-cell receptor¹⁹.

The Lyt 3-like chain of rat CD8 is distinct from the Lyt 2-like chain in that it also contains a sequence highly homologous to the joining (J) segments of immunoglobulin and the T-cell receptor, particularly those of light-chains. But in contrast, the J sequence of the rat CD8 gene does not rearrange and is encoded by the same exon as the V sequence.

The V sequences of both rat CD8 chains are more similar to light-chain V regions (~30%) than they are to each other (21%). Johnson and Williams suggest that this similarity to both V and J sequences implies that the CD8 genes have descended from genes encoding a heterodimer that existed before the development of gene-segment rearrangement. However, others have suggested that the gene encoding the Lyt 2 chain of CD8 actually is an orphan V-gene segment that lost its ability to rearrange¹. This view is supported by the extremely close linkage of human and mouse CD8 genes with immunoglobulin κ light-chain genes²⁰.

Domain structures of members of the immunoglobulin gene superfamily. *a*, Molecules of the immunoglobulin gene superfamily that are directly involved in antigen interaction (immunoglobulin, T-cell receptor and MHC class I and class II). The γ chain is included because of its structural similarity to the α and β chains. All these sequences are encoded by multigene families except β_2 -microglobulin, included here because of its interaction with the class I heavy chain. Immunoglobulin homology-unit sequences are represented by circles held by disulphide bonds. Only disulphide bridges present in homology units are indicated. *b*, Single-copy members of the family that are not directly involved in antigen interaction. CD4 and T3 δ have disulphide bridges indicated without homology-unit loops. These indicate sequences that may be descended from homology unit structures, but are now significantly divergent.



An intriguing alternative possibility for the evolution of the rat *Lyt* 3-like gene is that it is the product of an fortuitous *V-J* rearrangement and translocation in a germ cell. If correct, this suggests that the sequences and perhaps even enzymes responsible for the somatic rearrangement of immunoglobulin and T-cell receptor genes can also occasionally generate rearrangements in germ cells and consequently generate new gene combinations. Less similar *J*-like sequences have also been proposed for the *V* homology units of OX-2 and human CD4 (T4)^{16,8}, neither of which rearrange. The suggested *J* sequence of CD4 can be argued against by data that indicate it is split by an intron (J. Parnes, personal communication). Interestingly, mouse CD4 (L3T4) may contain another *J*-like sequence elsewhere in the molecule associated with the relic of another *V*-like region⁹. The *J*-like sequences in various single-copy members of the immunoglobulin gene superfamily may represent the product of different evolutionary pathways.

Though originally described as not belonging to the superfamily^{10,11}, re-examination of the δ - and ϵ -subunits of the human T-cell receptor-associated T3 molecule has led to the suggestion that both have a single *V*-like homology unit (A. Williams, personal communication). If these similarities prove convincing, it is striking that all the molecules directly associated with T-cell antigen and/or MHC recognition characterized to date (T-cell receptor, CD4, CD8 and T3) are or include members of the immunoglobulin gene superfamily.

A cDNA for α 1B-glycoprotein encodes five internally repeated domains most similar to *V* homology units¹⁴. Its structure is very like the poly-Ig receptor, which also has five tandem, distantly related, immunoglobulin-like domains. Although membrane-bound, poly-Ig has a secretory component and α 1B-glycoprotein may be the secretory component of a membrane-bound precursor¹⁴.

The most intriguing new member of the superfamily is a developmentally regulated neuronal cell-adhesion molecule, N-CAM. Its binding is homophilic (self) and probably polyvalent^{21,22}. The sequence of a partial cDNA isolated from chick embryo brain suggests that like poly-Ig and α 1B-glycoprotein, N-CAM has at least four homology-unit sequences, probably arising from internal duplication. Although described as *V*-like, these units are structurally more similar to *C* homology units. As many as 500 amino-terminal residues of N-CAM remain unknown, leaving the possibility that even more immunoglobulin homology units are present. The homophilic binding function of N-CAM has been mapped to the amino-terminal portion of the molecule along with the homology units²¹. Therefore the homo-

philic and polyvalent nature of N-CAM may result from receptor/ligand binding analogous to the paired homology unit associations that generate the domain structures of other molecules of the superfamily. If true, this finding supports models for the origin of the immunoglobulin superfamily that suggest primordial homology units possibly mediate cell-cell interactions through homophilic associations²³.

N-CAM is a member of a new functional class within the superfamily based on a polydomain structure. Others of this group include α 1B-glycoprotein and poly-Ig receptor; recent analyses suggest that even CD4 is descended from a polydomain precursor⁹. As the CD4 molecule is not known to form dimers, it may interact with class II molecules in a manner similar to the interactions of other polydomain molecules with their ligands, presumably in contrast to the heterodimeric interactions of the CD8 molecule. N-CAM is the first member of the immunoglobulin superfamily that has a brain-specific distribution.

A previously described brain protein, neurocytoplasmic protein 3, has recently been identified as a member of the superfamily (T. H., unpublished) and more recent extended analyses of this molecule suggest that it has four homology units (J.G. Sutcliffe, personal communication). Perhaps, like N-CAM, it functions in cell-cell interactions in the nervous system. Indeed, N-CAM demonstrates that members of the immunoglobulin superfamily can function outside the immune system, thus emphasizing the widespread distribution and versatility of functional charac-

teristics of its members.

Many additional members of the superfamily are surely yet to be identified—the intriguing question is how pervasive will be the usage of immunoglobulin homology units as recognition molecules in mammalian development. □

- Hood, L., Kronenberg, M. & Hunkapiller, T. *Cell* **40**, 225 (1985).
- Johnson, P. & Williams, A.F. *Nature* **323**, 74 (1986).
- Hill, R.L., Delaney, R. & Fellow, R.E. Jr *Proc. natn. Acad. Sci. U.S.A.* **56**, 1762 (1966).
- Amzel, L.M. & Poljak, R.J. *A. Rev. Biochem.* **48**, 961 (1979).
- Littman, D.R., Thomas, Y., Maddon, P.J., Chess, L. & Axel, R. *Cell* **40**, 237 (1985).
- Sukhatme, V.P., Sizer, K.C., Vollmer, A.C., Hunkapiller, T. & Parnes, J.R. *Cell* **40**, 591 (1985).
- Johnson, P., Gagnon, J., Barclay, A.N. & Williams, A.F. *EMBO J.* **4**, 2539 (1985).
- Maddon, P. *et al. Cell* **42**, 93 (1985).
- Tourville, B., Gorman, S.D., Field, E.H., Hunkapiller, T. & Parnes, J.R. *Science* (in the press).
- van der Elsen, P. *et al. Nature* **312**, 413 (1984).
- Gold, D.P. *Nature* **321**, 431 (1986).
- Mostov, K.E., Friedlander, M. & Blobel, G. *Nature* **308**, 37 (1984).
- Cunningham, B.A., Wang, J.L., Berggård, I. & Peterson, P.A. *Biochemistry* **12**, 4811 (1973).
- Ishioaka, N., Takahashi, N. & Putnam, F.W. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2363 (1986).
- Williams, A.F. & Gagnon, J. *Science* **216**, 696 (1982).
- Clark, M.J., Gagnon, J., Williams, A.F. & Barclay, A.N. *EMBO J.* **4**, 113 (1985).
- Hemperly, J.J., Murphy, B.A., Edelman, A.M. & Cunningham, B.A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3037 (1986).
- Sutcliffe, J.G., Milner, R.J., Shinnick, T.M. & Bloom, F.E. *Cell* **33**, 671 (1983).
- Goverman, J., Hunkapiller, T. & Hood, L. *Cell* **45**, 475 (1986).
- Sukhatme, V.P. *et al. J. exp. Med.* **161**, 429 (1985).
- Edelman, G. *Science* **219**, 450 (1983).
- Cunningham, B.A., Hoffman, S., Rutishauser, U., Hemperly, J.J. & Edelman, G.M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3116 (1983).
- Williams, A. *J. theor. Biol.* **98**, 221 (1980).

Tim Hunkapiller and Leroy Hood are in the Division of Biology 147-75, California Institute of Technology, Pasadena, California 91125, USA.

Chirality

Distinguishing true chirality from its accidental imitators

from Alexandra J. MacDermott

ONE of the strangest things about the universe is its handedness or chirality. A recent conference* assembled a variety of specialists, from particle physicists to psychologists, to discuss all aspects of chiral symmetry breaking. These ranged from the preponderance among spiral galaxies of rotation to the left with respect to the direction of recession, to the dominance of left cheeks in portraits. But the heart of the debate was the question of why life is based on DNA made of D-sugars and proteins made of L-amino acids, rather than the enantiomeric (mirror-image) L-sugar/D-amino-acid system.

Is the handedness of present-day biochemistry just a frozen accident? Or was it

determined by some chiral influence? Past candidates for the latter have included supposedly chiral combinations such as the Earth's gravitational and magnetic fields, or, more fancifully, asymmetric synthesis by stirring¹. According to L.D. Barron (Glasgow), however, these influences can now be discounted because they are 'falsely chiral'². The hallmark of 'true' chirality is natural optical rotation, which changes sign under space inversion but not under time reversal, that is, it is parity-odd, time-even. Absolute asymmetric synthesis can therefore only be induced by something with this same symmetry. Stirring (parity-even, time-odd) in the Earth's gravitational field (parity-odd, time-even) is therefore no good, as the overall influence is parity-odd, time-odd.

*Chiral symmetry breakings in physics, chemistry and biology, Rouen 17-19 June 1986.

Similarly, a magnetic field will not do, for it is parity-even (being an axial vector) and time-odd (imagine it generated by a circular motion of electrons, which changes direction under time reversal). Even in combination with the Earth's gravitational field, the overall influence remains falsely chiral (parity-odd, time-odd).

A good example of a truly chiral influence is a magnetic field (parity-even, time-odd) parallel to a light beam (parity-odd, time-odd), giving overall a parity-odd, time-even effect. The light beam need not be polarized. W. Thiemann (Bremen) reported the effect of this combination on the synthesis of hexahelicene. He obtained more L molecules with the field parallel to the light beam, and more D with the field anti-parallel. The enantiomeric excess, at 0.07 per cent, was just at the limits of detectability.

A deeper truly chiral influence comes from the weak interactions, mediated by the recently discovered W^- and Z^0 bosons (ref. 3). The weak force gives an intrinsic left-handedness, or helicity, to the electron: radioactive β -decay produces an excess of electrons with anti-parallel (left-handed) momentum and spin vectors over the alternative parallel (right-handed) combination, which increases in proportion to v/c , the velocity of the electrons relative to the velocity of light. Similarly positrons emitted in β -decay have an intrinsic right-handedness. Left and right-handed molecules are therefore not true enantiomers, because the weak interaction can distinguish them. The true enantiomer of an L-amino acid is the D-amino acid made of anti-matter. This parity-violating effect of the weak interaction is detectable in a very slight optical activity of atoms, measured by P.G.H. Sandars (Oxford) for Bi vapour⁴. His result agrees with theory to within 20 per cent, and if this could be improved to just a few per cent, different unified field theories could be distinguished, giving us 'table-top' particle physics. J.B. Robert (Grenoble) hopes to detect a slight difference in the nuclear magnetic resonance chemical shifts of enantiomers.

Weak interactions

The symmetry-breaking effects of the weak interactions are very small, but could be amplified to produce today's handed biochemistry. The largest effect is the differential radiolysis of left- and right-handed molecules by left-helical electrons from β -decay. This was discussed by A.S. Garay (College Station, Texas), R.A. Hegstrom (Wake Forest, North Carolina) and A. Rich (Michigan), who predicted an asymmetry (ratio of enantiomeric excess to total number of molecules) of about 10^{-10} . P.S. Farago (Edinburgh) reported a beautiful experiment⁶ on the differential absorption by chiral molecules of left and right helically polarized 5-eV electrons,

which gave an unexpectedly large asymmetry of 10^{-4} with an electron beam of 0.5 per cent excess left helicity. But many participants saw problems in the deceleration of β -decay electrons to energies gentle enough not to blow the molecules apart. This is done very carefully in laboratory experiments, but in nature it involves scattering, which could spoil the helicity.

Circularly polarized photons could cause a similar chiral selection, but without the need for a nearby radioactive source. Unfortunately the slight circular polarization of sunlight around sunrise is balanced by an equal and opposite polarization around sunset, so this mechanism would only be viable if life evolved in a lake bounded to the east or west by a large mountain.

All the above mechanisms rely on specific local conditions, and the enantiomer selected depends on time and place. A global mechanism was discussed by S.F. Mason (King's College, London) and G.E. Tranter (Oxford), who point out that since left- and right-handed molecules are not true enantiomers, they differ slightly in energy. This parity-violating energy difference (PVED), arising from weak neutral current interactions between the electrons and nuclei, was calculated⁷ by Tranter for several amino-acids, and in each case the L-amino acid was the more stable, as were the L-residues in both the α -helix and β -sheet polypeptide conformations. The asymmetry factor, at 10^{-17} , corresponding to an excess of 10^6 L-amino acid molecules per mole, is smaller than that for β -radiolysis, but applies to all chiral molecules at all times. Tranter found equal success with the sugar series: D-glyceraldehyde, a prebiotic precursor of more complex sugars, is more stable than its L-enantiomer, as is D-ribose in the C₂-endo conformation (that found in DNA).

These small asymmetries need amplifying. D.K. Kondepudi (Austin, Texas) has devised a kinetic scheme⁸ whose essential feature is autocatalysis, that is, the presence of one enantiomer aids its production. With concentrations as low as 10^{-3} molar, and amino-acids being produced at, say, 1.5 moles per litre per 100 years, the initial PVED asymmetry of 10^{-17} can be amplified to 98 per cent homochirality in just 10^4 years. S.L. Miller (San Diego) argued that racemization would cause difficulty. Of course 'racemic' mixtures are not quite equimolar because of the PVED, so a problem exists only if racemization is fast compared with the amplification timescale. In fact Kondepudi's mechanism includes racemization, and can withstand a racemization half-life as low as 260 years, compared with the minimum of 10^4 years reported for real molecules. But controversy continues as to whether Kondepudi's assumed production rate is too high. The greater initial

asymmetry from β -radiolysis could, of course, be amplified in a much shorter time, with lower minimum production rates and concentrations.

Ideal candidates for the Kondepudi amplification mechanism are polymerization reactions: several speakers described experiments showing that these proceed best in homochiral monomer solutions, and tend to terminate if an oppositely handed monomer adds to the chain.

Giant molecules

Giant chiral molecules, such as polymers and especially crystals, may have greatly enhanced PVED, which Mason and Tranter believe accounts for the consistent 1 per cent excess of L-quartz crystals in samples from all over the world. The role of mineral catalysts in prebiotic chemistry has aroused great interest: L-quartz, for example, preferentially absorbs L-alanine from a racemic mixture for purely diastereomeric reasons, with an enantio-selectivity of 1 per cent. Combined with the 1 per cent excess of L-quartz from the PVED, prebiotic surface catalysis, according to Tranter⁹, could provide an asymmetry factor of 10^{-4} , which is much greater than the 10^{-17} calculated for individual molecules, and could explain today's homochiral biochemistry with a much less powerful amplification than Kondepudi's.

Another central question is whether chiral purity is a pre-condition or a consequence of life. It is generally accepted as a precondition, as it is needed for efficient polymerization. But Miller disagrees, believing that chiral selection can occur only biologically, not chemically. He described a glycerol-based polymer which could form a non-chiral replicator: glycerol, he says, is easy to make, but ribose is difficult, and "if it isn't easy it's probably not prebiotic".

An excellent laboratory in which to test for chiral selection during purely chemical evolution is the interesting atmosphere of Titan, and A. Brack (Orleans) suggests that the forthcoming Cassini mission should carry out optical-rotation measurements. Finding the same handedness in different extra-terrestrial systems would add support to the weak interaction as a global symmetry breaker. □

1. Dougherty, R.C. *J. Am. Chem. Soc.* **102**, 380 (1980).
2. Barron, L.D. *Chem. Soc. Rev.* **17**, 189 (1980).
3. Walgate, R. *Nature* **303**, 473 (1983).
4. Sandars, P.G.H. in *Fundamental Interactions and Structure of Matter* (eds Crave, K., Duclos, J., Fiorentini, G. & Torelli, G. 57 (Plenum, New York, 1980).
5. Hegstrom, G.A. *Nature* **297**, 643 (1982).
6. Campbell, D.M. & Farago, P.S. *Nature* **318**, 52 (1985).
7. Tranter, G.E. *Mol. Phys.* **56**, 825 (1985).
8. Kondepudi, D.K. & Nelson, G.W. *Phys. Rev. Lett.* **50**, 1023 (1983).
9. Tranter, G.E. *Nature* **318**, 172 (1985).

Alexandra J. MacDermott is at the Physical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK.

Physiology

Two glucagon transducing systems

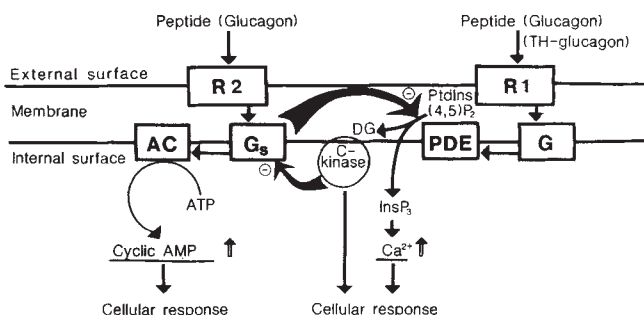
from O.H. Petersen and C. Bear

THE classic studies by Sutherland and co-workers¹ on the mechanism of action of glucagon and adrenaline on the liver are fundamental to the concept that many hormone actions are mediated by the intracellular messenger cyclic (c)AMP. Although it is now clear that there are other important intracellular messengers^{2,3}, for example, inositol trisphosphate, glucagon remains a textbook example of a hormone acting via cAMP (ref. 4). The report by Wakelam *et al.* on page 68 of this issue⁵ challenges this view by demonstrating that liver cells possess two distinct receptors for the peptide hormone glucagon, one type coupled to a phosphodiesterase, causing breakdown of phosphatidylinositol-4,5-bisphosphate, and the other coupled to adenylate cyclase, generating cAMP (see figure).

The grounds for this discovery were laid by Silvia Corvera and colleagues⁶ who used a glucagon analogue made by trinitrophenylation of the α -NH₂ group of 12-homoarginine glucagon (TH-glucagon). TH-glucagon is a full agonist for the stimulation of glycogenolysis, gluconeogenesis and urea synthesis in hepatocytes, but is only a very weak partial agonist for cAMP accumulation⁶.

Wakelam *et al.*⁵ now show that both TH-glucagon and glucagon cause a rapid dose-dependent increase in the production of inositol 1,4,5-trisphosphate (InsP₃). Glucagon-stimulated production of InsP₃ is biphasic; inhibition occurs at higher glucagon concentrations when adenylate cyclase activity is maximally stimulated. No such inhibitory phase is observed with TH-glucagon, suggesting that a component of the cAMP-linked transduction system inhibits the production of inositol phosphates. Indeed, TH-

glucagon does not stimulate adenylate cyclase whereas glucagon causes half-maximal stimulation of this enzyme at a concentration of about 6×10^{-9} M. In contrast, about 3×10^{-10} M of both glucagon and TH-glucagon cause half-maximal stimulation of inositol phosphate production. At physiological glucagon concentrations ($\leq 10^{-10}$ M) activation of the adenylate



Schematic diagram of the two signal-transduction systems. R, receptor; G, GTP-binding protein; AC, adenylate cyclase; PDE, phosphodiesterase; DG, diacylglycerol; PtdIns(4,5)P₂, phosphatidylinositol-4,5-bisphosphate.

cyclase-cAMP pathway is less significant than the main effect mediated via the phosphodiesterase-diacylglycerol-InsP₃ pathway. From these new data⁵ it seems that glucagon is very far from being a typical hormone acting solely via cAMP, but instead is a good example of a hormone that uses calcium ions and diacylglycerol as intracellular messengers.

There is an analogy between the proposed dual-receptor system for glucagon and the two receptor mechanisms of vasopressin^{4,7}, but the two types of vasopressin receptors are not both present in the same cells. The V₁ (inositol phosphate-linked) receptor for vasopressin is located on liver cells whereas the V₂ (cAMP-associated)

receptors are found on renal cells.

It is unlikely, however, that the two separate receptor systems proposed for glucagon, which co-exist in the liver cells, are unique to this hormone. It has very recently been shown by Trimble *et al.*⁸ that the peptide secretin (which is chemically related to glucagon) can activate the two separate second-messenger systems in pancreatic acinar cells. Cholecystokinin is now recognized as a peptide hormone activating the Ca²⁺ pathway⁹, but it was demonstrated many years ago that it can activate adenylate cyclase in broken cell membrane preparations¹⁰, although in intact pancreatic acinar cells no cAMP accumulation occurs⁹.

This old, unexplained finding may relate to the recent results of Wakelam *et al.*⁵ in suggesting that protein kinase C activation evokes desensitization of glucagon-stimulated adenylate cyclase. Conversely it now seems that activation of adenylate cyclase can inhibit inositol phosphate production, possibly mediated via the stimulatory guanine

nucleotide regulatory protein G_s (ref. 5; see figure).

The newly discovered complexity of the action of glucagon is an example of how difficult it is to elucidate the mechanisms underlying hormonal effects. In the case of glucagon it had been firmly established that the metabolic actions¹ as well as the immediate effects on the membrane potential¹¹ can be mimicked by the application of exogenous dibutyl cAMP, and there is abundant evidence that it stimulates adenylate cyclase and produces cAMP (ref. 1).

All these results still stand, but it is now clear that there is a second, physiologically more important messenger pathway, although the complex and puzzling interactions between the two signal-transduction systems will need much more work before they are finally sorted out. □

1. Robison, G.A. *et al.* *Cyclic AMP* (Academic, New York, 1971).
2. Nishizuka, Y. *Nature* **308**, 693 (1984).
3. Berridge, M.J. & Irvine, R.F. *Nature* **312**, 315 (1984).
4. Berridge, M.J. *Scient. Am.* **253** (4), 142 (1985).
5. Wakelam, M.J.O. *et al.* *Nature* **323**, 68 (1986).
6. Corvera, S. *et al.* *Biochim. biophys. Acta* **804**, 434 (1984).
7. Creba, J.A. *et al.* *Biochem. J.* **212**, 733 (1983).
8. Trimble, E. *et al.* *Biochem. J.* (in the press).
9. Schulz, I. & Stolze, H.H. *A. Rev. Physiol.* **42**, 127 (1980).
10. Ritten, W.J. *et al.* *Biochim. biophys. Acta* **274**, 201 (1972).
11. Petersen, O.H. *J. Physiol., Lond.* **239**, 647 (1974).

O.H. Petersen, George Holt Professor of Physiology, and C. Bear, a visiting Research Fellow from the University of Calgary, Canada are in the Department of Physiology, University of Liverpool, Liverpool L69 3BX, UK.

Molecules link bats in New Zealand and South America

Mystacina tuberculata, the lesser short-tailed bat, was long thought to be one of New Zealand's archaic, mystery vertebrates. But the paper by Elizabeth D. Pierson and colleagues on page 60 of this issue describes immunological comparisons suggesting that the closest phylogenetic affinity of this species is with the New World phyllostomids. This specimen is on the bark of a southern rata and its unique features — mole-like fur, robust hind limbs and tightly folded wings, all adaptations for its terrestrial habits, are clearly visible. Photograph by Elizabeth Pierson of the Museum of Vertebrate Zoology, University of California, Berkeley, California 94720, USA.



History of science

Isotope geochemistry and cosmochemistry

from R. Letolle

ONE hundred years ago, Sir William Crookes (1832–1919) explained at the British Association meeting in Birmingham some of the principles then relevant to problems in geoscience and nucleosynthesis, and argued for the existence of isotopes. Twenty-four years before the conceptualization of isotopes by Soddy, Crookes published the results of his reflections (*Nature* 34, 423; 2 September 1886), showing an extraordinary presence of some present-day concepts of isotope chemistry and geochemistry.

In 1911, Soddy had just discovered yttrium, and the intricacy of spectroscopic lines of the mixture of rare-earth elements led him to conclude that elements could be a mixture of some 'meta-elements'. He recalled Prout's system when he wrote:

"... Let us picture the very beginnings of time, before geological ages, before the Earth was thrown from the central nucleus of the original *protyle*. Let us still imagine that at this primal stage all was in an ultra-gaseous state, at a temperature inconceivably hotter than anything now existing in the visible Universe; so high, indeed, that the chemical atoms could not yet have been formed, being still far above their dissociation point..."

From Crookes' Birmingham address, 24 years earlier:

"We require a word, analogous to protoplasm to express the idea of the origin primal matter existing before the evolution of the chemical elements. The word I have ventured to use for this purpose is compounded of $\pi\rho\sigma$ (earlier than) and $\upsilon\lambda\eta$ (the stuff of which things are made). The word is scarcely a new coinage, for 600 years ago Roger Bacon wrote in his *De Arte Chymiae*: 'The elements are made out of $\upsilon\lambda\eta$, and every element is converted into the nature of another element'."

Relative to the possibility of unstable atoms and the chemical fractionation of atoms of various masses of the same element, Crookes says:

"... I have said that the original *protyle* contained within itself the potentiality of all possible atomic weights. It may well be questioned whether there is an absolute uniformity in the mass of every ultimate atom of the same chemical element. Probably our atomic weights merely represent a mean value around which the actual atomic weights of the atoms vary within certain narrow limits.

Each well-defined element represents a platform of stability connected by ladders of unstable bodies. In the first accreting together of the primitive stuff the smallest atoms would

form, then these would join together to form larger groups, the gulf across from one stage to another would be gradually bridged over, and the stable element appropriate to that stage would absorb, as it were, the unstable rungs of the ladder which led up to it. I conceive, therefore, that when we say the atomic weight of, for instance, calcium is 40, we really express the fact that, while the majority of calcium atoms have an actual weight of 40, there are not a few which are represented by 39 or 41, a less

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

William Crookes

number by 38 or 42, and so on. We are here reminded of Newton's 'old worn particles'.

It is not possible, or even feasible, that these heavier and lighter atoms may have been in some cases subsequently sorted out by a process resembling chemical fractionation? This sorting out may have taken place in part while atomic matter was condensing from the primal state of intense ignition, but also it may have been partly effected in geological ages by successive solutions and re-precipitations of the various earths.

This may seem an audacious speculation, but I do not think it is behind the power of chemistry to test its feasibility..."

After that Crookes presents his own experimental evidence of the reality of various species of yttrium atoms, based on spectroscopic data. In fact, yttrium is monoisotopic, and isotope effects in radiation spectra were discovered experimentally much later. Nevertheless, Crookes' idea is true and he in fact observed the complex array of spectral lines resulting from various rare-earth elements and some of their combinations. Some of his conclusions merit quotation:

"Returning ... to the idea of heavy and light atoms, we see how well this hypothesis accords with the new facts here brought to light. From every chemical point of view the stable molecular group, yttrium, behaves as an element.

Excessive and systematic fractionation has acted the part of a chemical 'sorting Demon', distributing the atoms of yttrium into several groups, with certainly different phosphorescent spectra, and presumably different atomic weights, though all these groups behave alike from the usual chemical point of view. Here, then, is one of the elements the spectrum of which does not emanate equally from all its atoms, but some atoms furnish some, other atoms others, of the lines and bands of the compound spectrum of the element. And as this is the case with one element, it is probably so in a greater or less degree with all. Hence the atoms of this element differ probably in weight, certainly in the internal motions they undergo.

Another important inference which may be drawn from the facts is, that the atoms of which yttrium consists, though differing, do not differ continuously, but *per saltum*. We have evidence of this in the fact that the spectroscopic bands characteristic of each group are distinct from those of the other groups, and do not pass gradually into them. We must accordingly expect, in the present state of science, that this is probably the case with the other elements. And the atoms of a chemical element being known to differ in one respect may differ in other respects, and presumably do somewhat differ in mass.

Restricted by limited time and means, even a partial separation of these atomic groupings is possible only with enormous difficulty..."

Coming back to the origin of elements, Crookes discusses at length a theory of element-building from *protyle* (note that the word $\upsilon\lambda\eta$ was introduced again by Gamow in the 1940s in his theory of the Big Bang) and wonders about the gap between uranium and thorium from lighter elements (in lead). He concludes:

"... What comes after uranium? I should consider that there is little prospect of the existence of an element much lower than this..."

The explanation of the absence of elements between thorium and lead, which refers to the instability of such elements, is ingenious and, in the context of knowledge at that time, is not far from the truth.

There are in this wonderful and forgotten paper many striking ideas on other problems, for instance:

"... A genesis of the elements such as is sketched out would not be confined to our little Solar System, but would probably follow the same general sequence of events in every centre of energy now visible as a star."

and:

"[it is] ... a question that I especially commend to the younger generation of chemists, not only as the most interesting, but the most profoundly important, in the entire compass of our science." □

R. Letolle is in the Department of Geological Dynamics, Université Pierre et Marie Curie, 75252 Paris Cedex 05, France.

Minimum-Coulomb-energy electrostatic configurations

SIR—Determination of the stable (least energy) configuration of N equal charges on the surface of a unit sphere remains incomplete¹ despite the importance of the problem in botany, nuclear theory and elsewhere^{1,2}. Recently Munera² presented an elegant description of three-dimensional configurations in which the charges are partitioned among the vertices of regular polygons, parallel to the equatorial plane of the sphere. Munera labels the results as possessing minimum Coulomb energy. But configurations with lower energies can be found. Munera's Table 2 should therefore be more clearly labelled as the minimum energies obtained by numerical optimization subject to the strict (and over-limiting) geometrical constraints applied. Indeed Munera seems to acknowledge the existence of a lower energy distribution for $N=11$ and also states that "rotational equilibrium does not imply minimum energy". Here it is indeed true that the solutions presented improve on several previously reported, and conveniently involve vertices of polyhedra with triangular faces, but they are nevertheless suboptimal and derivative physical interpretations are inaccurate.

Truly optimal configurations have been obtained by numerical Monte Carlo techniques in which no constraints on geometric regularity are applied. Whilst one must accept that the results of such a technique only generate true minima for infinite computing time, one cannot dispute that the configurations generated by finite computations are more stable and lower energy configurations than those calculated by Munera have been found for $N=11, 13, 15, 19, 20$. As an example a configuration with $W(13)=58.8535$ (origin=sphere centre) is given below:

x	y	z
0.00000	0.00000	1.00000
-0.22004	-0.97369	-0.05925
-0.79907	-0.31927	0.50946
0.78227	0.38408	0.49045
0.21185	-0.51330	-0.83165
-0.20457	0.40678	-0.89033
0.87263	-0.48031	-0.08844
-0.30184	0.73503	0.60715
0.75763	0.33346	-0.56107
0.18079	0.97225	-0.14849
0.34187	-0.71106	0.61443
-0.77193	-0.35116	-0.52992
-0.84741	0.51760	-0.11829

Removal of the symmetry constraint denies a simple geometrical shorthand but the result cannot be ignored.

Munera's demonstration that the effective equipotential surface resulting from a distribution of finite charges can lead to classical transparency to a non-relativistic particle is exciting in view of a novel way of interpreting nuclear scattering and radioactive decay. The prediction of a finite number of directions in which

'holes' in the equipotential surface arise may, however, be a due to the assumed symmetries and before further physical interpretations it is important to establish whether the same phenomena arise when the constraint is removed as for example with the numerical data reported here.

S WEBB

Joint Section of Physics,
Institute of Cancer Research
and Royal Marsden Hospital,
Sutton, Surrey SM2 5PT, UK

1. Berezin, A.A. *Chem. Phys. Lett.* **123**, (1,2), 62-64 (1986).
2. Munera, H.A. *Nature* **320**, 597-600 (1986).

Haldane's rule back in the news

SIR—There has been a revival of interest in Haldane's rule^{1,2}; that "When in the F_1 offspring of two different animal races one sex is absent, rare or sterile, that sex is the heterozygous (heterogametic) sex"³. Haldane felt that if his rule "is more than a mere coincidence" it must show how lethality, sterility and, in some instances, the sex transformation of the heterogametic sex in hybrids, "may be explained as due to the same cause"³. In trying to understand the genetic basis of Haldane's rule, it might be illuminating to view such sex-specific fitness effects (in the offspring of interspecific hybrids) in the larger context of sex-specific differences related to mechanisms of sex determination.

Haldane preferred not to accept the obvious explanation in terms of sex factors because clear-cut examples of such genes were unknown. He sought an explanation in terms of autosomal genes, a few of which are known to cause sex-specific effects on fitness. Properties of genes such as *daughterless* (*da*) and *male-lethal* (*mle*) in *Drosophila melanogaster* take on added significance in the light of Haldane's views. These genes are autosomal but they are part of the sex-determination mechanism and appear to act to cause sex-specific lethality^{4,6}.

Haldane maintained that heterogameity rather than sex is the critical feature responsible for the sex-specific fitness effects observed in interspecific hybrids³. It may therefore be of interest to note that according to a model developed by us to understand X/A ratio measurement and the effects of *da*, *mle* and certain mutations at the X-linked locus *sex-lethal*, the narrow range of viability of the male *D. melanogaster* is a consequence of his heterogametic condition⁶.

It appears that sex transformation of the heterogametic hybrid — more than its lethality or sterility — requires explanation in terms of incompatibility between the sex-determining mechanisms of the parental species or races. A case of sex reversal in hybrid mice⁷ has been inter-

preted on this basis⁸. Data from these hybrid mice, in particular differences between the F_1 and backcross progeny in the frequency and degree of sex transformation, may be explained as being due to genetic interaction among an autosomal gene, an X-linked gene and the Y chromosome⁸.

H. SHARAT CHANDRA

MCBL & ICMR Centre,
Indian Institute of Science,
Bangalore 560 012,
India

1. Coyne, J.A. *Nature* **314**, 736-738 (1985).
2. Licht, L.E. & Bogart, J.P. *Nature* **317**, 207 (1985).
3. Haldane, J.B.S. *J. Genet.* **12**, 101-109 (1922).
4. Cline, T.W. *Genetics* **90**, 683-698 (1978).
5. Belote, J.A. & Lucchesi, J.C. *Nature* **285**, 573-575 (1980).
6. Gadagkar, R., Nanjundiah, V., Joshi, N.V. & Chandra, H.S. *J. Biosci.* **4**, 393-415 (1982); *Curr. Sci.* **50**, 805-807 (1981).
7. Eicher, E.M., Washburn, L.L., Whitney, J.B. & Morrow, K.E. *Science* **217**, 535-537 (1982).
8. Chandra, H.S. *Molec. gen. Genet.* **192**, 384-388 (1984).

Radiation and fetal brain development

SIR—Your article about the reactor accident at Chernobyl¹ ignores the fact that radiation sickness, cancer and mutations are not the only important effects of excessive radioactivity.

Recent studies in our laboratory²⁻⁴ and reports from others have revealed that radiation exposure of pregnant animals produces permanent brain deficiencies in their offspring, even at doses too low to produce radiation sickness. Early exposures are particularly damaging, because radiation affects DNA replication in the early proliferation of neurones. Since, in humans, neurone proliferation terminates before the eighteenth week of pregnancy radiation exposure may be particularly harmful at a time when a woman may not yet know that she is pregnant. Thus an evacuation of women known to be pregnant (as happened at Chernobyl), even if it can be accomplished immediately, would still leave those women unaware of their pregnancy at risk of fetal brain damage.

It is important to note that, in contrast to cancer and genetic damage (mutations), which are rare events, brain damage could occur in almost every fetus exposed to radiation.

STEPHEN ZAMENHOF

Mental Retardation Research Center
and Brain Research Institute,
UCLA School of Medicine,
Los Angeles, California 90024,
USA

1. *Nature* **321**, 100 (1986).
2. Zamenhof, S. & van Marthens, E. *Radiat. Res.* **77**, 117-127 (1979).
3. Zamenhof, S. & van Marthens, E. *Radiat. Res.* **85**, 292-301 (1981).
4. van Marthens, E. & Zamenhof, S. in *Developmental Effects of Prenatal Irradiation* (eds Kriegel, H., Schmal, W., Kistner, G. & Stieve, F.E.) 92-103 (Gustav Fisher, New York, 1982).

Maestros and bureaucrats

Duncan Davies

The Tender Ship: Governmental Management of Technological Change.

By Arthur M. Squires. Birkhauser: 1986. Pp.247. \$24.95.

BUREAUCRACIES that disburse money selectively are feared and disliked by the recipients, even when the rules are clear and the process just. When, as in the funding of science and technology, the basis for choice is difficult or even impossible to codify, the dissatisfaction is permanent and acute. It is therefore a great achievement to write on this subject so as to fascinate, delight and inform the reader; Dr Squires has done just this. His secret is to tell a large number of case stories, ranging over centuries, using his broad perceptions as a teacher, a government contractor and helper, and a multidisciplinary researcher. So horribly and entertainingly does he expose the inadequacies of the present system, that radical reform seems essential.

Cleverly chosen material has been assembled by fourth-year students for seminars. In considerable detail, it shows how items as varied as a 17th-century Swedish ship, the airship R101, a putative nuclear aeroplane, water desalination systems and an army rifle acquired unsuitable and idiotic design requirements and timetables as a result of personal ambition, fear, politicking, incompetence and place-seeking by remote officials. Contrasting and successful cases include the development of a gaseous diffusion plant to separate the uranium isotopes (in which project Squires was involved as a young worker), radar in Britain and the American wartime programme to develop synthetic rubber.

The key feature the successes share seems to be direction by "maestros of technology", who have large networks of equally competent friends to be brought in at critical moments. Another difference is that the successes began as essential wartime needs, based on clear decisions. The failures were peacetime projects lacking immediate and urgent need, which made them victims of entangled manoeuvring about design and priority. The nuclear aeroplane project, for example, oscillated between preference for direct heating of the working fluid, to indirect heating via a transfer fluid, so that a very doubtful starter never made headway. By contrast, those determining the course of the gas diffusion plant for uranium hexafluoride were "maestros of technology" motivated entirely by the urgent need to succeed. The one arbitrarily imposed change was to abandon flat membranes in favour of tubes. It was traumatic, and discarded

much work, but it made sense to all those who were involved, and accelerated progress.

Squires believes in the mixed economy: the private sector for the product needed on a short-to-medium timespan for a price that definite customers will pay; and the public sector for the basic work, the long-range job and items that society needs but customers do not buy. It is interesting that research for the marketplace lies between his extremes. Target specifications for a vehicle, domestic gadget, telecommunications feature or food product are bound to change as the market alters, and become better understood as prototypes are tried out. Although such changes in aim are infuriating, expensive and disheartening, the market can be seen and understood better than a distant bureaucracy. It is known who is being captious or demanding, and to some extent, why. Utility, attractiveness and cost-effectiveness are not as incomprehensible and irrational as politicised fund-control. This adds point to the author's discontent about the funding methods used by governments in peacetime.

As might be expected from such an active and lively mind, the solution cautiously proposed is draconian and revolutionary. For the funding of basic science, Squires suggests replacing the incompetent and sometimes dishonest decision of the remote bureaucrat or committee by the judgement of applicants' previous research results rather than prospectuses,

and then paying 80 per cent of the prize money as block grants to their institution, with only 20 per cent going to the individual. Selection of the research projects would then be a local matter, and encourage laboratories to bid hard for the best people. Such an arresting idea is certainly bound to stimulate radical discussion.

The assessments for prize-awarding on this scale, however, would be very demanding on the time of able scientists, and it is hard to see that orthodoxy would not prevail. One could hardly find enough competent juries comprising people not doing research; peer-judgement seems inevitable, so that this week's judge will be next week's appellant. More thought seems to be needed, to avoid jumping from an uncomfortably hot frying pan into a red-hot fire.

Alongside the descriptions of work, there are perceptive analyses of the habits of bureaucracies, presidents, practitioners and professors, on whom Squires' distilled wisdom is well worth close attention. There are some mistakes; it seems unjust to say that the Solvay process for making alkali "quickly displaced the Leblanc process everywhere except in England" (p.119) when it was a British firm that built the first large Solvay plants and then exported them widely. However, with so much material so attractively assembled such errors are almost inevitable. After the first reading, this is a book for the bedside, to be dipped into from time to time. □

Duncan Davies, 3 Broadlands Close, London N6 4AF, UK, formerly Chief Scientist and Engineer, Department of Trade and Industry, and Research and Development General Manager of ICI, is Chairman of British Ceramics Research Association Ltd and President of the Society of Chemical Industry.

Molecular patterns

Robin Carrell

The New Genetics and Clinical Practice, 2nd Edn. By D.J. Weatherall. Oxford University Press: 1986. Pp.206. Hbk £24.95, \$15.95; pbk £8.95.

THE CHANGE in human genetics from a descriptive to a molecular science has occurred so rapidly that the discipline has been left with an outdated terminology and outmoded concepts. The consequence is a need for a profound revision of the whole subject to give what has become known as the "new genetics". It is not surprising that the clearest exponents of this new genetics have approached it indirectly, not as traditional geneticists,

but laterally through the molecular studies of individual inherited diseases. This is the case with Professor Weatherall whose background has been the investigation of the disorders of haemoglobin synthesis.

The abnormal haemoglobin variants have long demonstrated that genetic diseases are usually syndromes with the same disease picture arising from a number of diverse mutations. Over 100 different mutations are known to give the characteristic anaemia of the unstable haemoglobins, of which nearly a third arise from new spontaneous mutations. This is the pattern now being seen in haemophilia where again a well-defined disease arises from a whole range of molecular variations, the occurrences of the disease being sustained by the continuous addition of new mutations.

As well as the sporadic mutations there are also the well-conserved polymorphisms of the globin genes, such as haemoglobin S and the thalassaemias. Here each mutation conveys an advantage to the heterozygote, as a defence against malaria, that balances the inherent lethal disadvantage to the homozygote. Again the surprising feature is the diversity of lesions, as illustrated by the identification of 50 different modes of inheritance of just one of the thalassaemia syndromes, haemoglobin H disease. Even haemoglobin S disease, which is due to a single amino acid substitution, is now known to have arisen independently in two or more geographical locations. Again the pattern seen with the haemoglobinopathies is also being observed in more subtle diseases of other proteins. Alpha-1-antitrypsin deficiency has arisen in the European from two independent mutations, both of which give a similar phenotype and some, as yet undetermined, past advantage which has ensured their survival. No doubt the overall pattern will repeat itself again with the other polymorphisms that give common inherited diseases such as phenylketonuria and cystic fibrosis.

This diversity of lesions means that any comprehensive text of human genetics is likely to resemble a catalogue. Weatherall avoids this problem by confining his account to the principles of the new genetics. His book, a complete rewrite of what was originally a small monograph conveys the excitement, potential and problems of the "new" science. It provides a clear, readable account that spans basic molecular genetics and the molecular pathology of inherited diseases, as well as the human and moral issues involved. This has all been achieved in some 200 pages by a well-planned coordination of the text with more than 50 clearly labelled, two-colour line diagrams.

A strength of the book is its concise descriptions of the technical terminology, old and new, both in molecular and chromosomal genetics. It is a volume that should be widely used in all stages of medicine and will be a useful additional text in general biology from advanced school level onwards. Its vitality should inspire students with the prospects that lie ahead and help overcome the lethargy in molecular medicine that Professor Weatherall has quite correctly diagnosed in Britain. □

Robin Carrell has recently taken up the Professorship of Haematology at the University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK.

Also recently published is David Galton's *Molecular Genetics of Common Metabolic Disease*, which is more modest in its ambitions but will be valuable as a briefer introduction to molecular genetics, particularly for the trainee physician or pathologist. Published by Edward Arnold. Pbk £10.50.

The next cold war?

Francis Auburn

Minerals and Mining in Antarctica: Science and Technology, Economics and Politics. By Maarten J. de Wit. Clarendon: 1986. Pp.127. £22.50, \$29.95.

The Seventh Continent: Antarctica in a Resource Age. By Deborah Shapley. *Resources for the Future*, 1616 P Street NW, Washington DC 20036:1986. Pp.315. \$40, £28.50.

ANTARCTICA is a unique continent governed in a unique manner. Under the Antarctic Treaty, which came into force in 1961, the nations claiming territories and those disputing claims have agreed not to pursue their rivalries whilst the Treaty is in force. The continent is devoted to peaceful purposes. Military activities, nuclear explosions and the disposal of radioactive wastes are forbidden, and all areas are open to inspection by designated observers at all times.

The Treaty regime, operating through meetings of an inner circle of Consultative Parties from 18 nations, has until recently been unchallenged. In good part this was due to the fact that the predominant activity in the area was fundamental scientific research, which served the dual purpose of enabling governments to maintain stations, and therefore a political presence, whilst at the same time lessening the potential for conflict. As long as the area was perceived as being of minor political interest on the world stage, there was little incentive to disturb the *status quo*.

Since 1970 a number of factors have conspired to raise the stakes. The sharp increases in the price of oil after 1970 put pressure on oil companies to find new reserves, and geological analogies suggest that the Antarctic continental shelf must contain substantial oil deposits. Because the Treaty was drafted to govern scientific research, it is unsuitable for petroleum licensing and the Consultative Parties slowly came to realize that a minerals regime is needed in advance of resource exploration. However it is only recently that negotiations, in the form of a regularly convened Special Consultative Meeting, have got down to the practical and difficult issues involved.

There is also the question of fisheries. In the mid-1970s the expansion of 200-mile exclusive zones forced the principal offshore fishing nations to seek new stocks. Krill, a small shrimp-like crustacean of the Southern Ocean, had the potential for becoming the world's largest fishery if technological and economic problems could be overcome. In 1980 the Consultative Parties negotiated the Convention on the Conservation of Antarctic

Marine Living Resources, setting up a Commission to conserve and regulate the Antarctic marine ecosystem. In 1983, following the drafting of the Law of the Sea Convention, which defined the seabed as the common heritage of all mankind, developing countries turned to the UN General Assembly in an effort to have the same concept applied to Antarctica.

The books by De Wit and by Shapley focus on these resources issues which have gradually overtaken science as the focus of attention in Antarctica. The volumes are a contribution to the ongoing debate about the future of the continent.

The minerals negotiations have proceeded on the assumption that Antarctica's onshore minerals will certainly not be exploited for several decades — and probably never — because of the extent of ice-cover (98 per cent), the continent's remoteness, the lack of appropriate technology and the availability of such minerals elsewhere. The central thesis of De Wit's book is that this view may be mistaken "groupthink", and in particular that the platinum-group metals of the Dufek Massif, close to the Weddell Sea in the sector south of Argentina and Chile, are likely to be economically viable at current prices. De Wit proposes an exploration programme at a relatively modest cost (up to \$10,000,000). These views, cogently argued by a geologist who has worked in the Antarctic, have serious implications for the minerals negotiations.

De Wit stresses the predominance of South African production and exports to the Western world of platinum-group metals, and the unreliability of this source given that country's current problems. (The political ramifications of apartheid have now spread further. After the book was written, the General Assembly of the United Nations resolved to urge the Consultative Parties to exclude South Africa from Antarctic decision-making.) In emphasizing the strategic role of platinum-group metals, the author indicates the need for the Western world to find alternative sources. In all this, there is an interesting parallel with the decision of the United States to allow chrome imports from the Smith regime of southern Rhodesia.

Much of the research on the Dufek site has been carried out by the United States, but the Soviet Union also has a clearly expressed interest in the area and its minerals. De Wit points out that the Dufek Massif is in the area claimed by Argentina, Chile and the United Kingdom, and that their programmes too have recently placed greater emphasis on resources. His thesis therefore raises the possibility that economic exploration for strategic onshore minerals is being carried out in an area of interest to three claimants and the superpowers. A theme of the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Antarctica: the minerals issue could undermine compromise.

Consultative Parties' defence of the Antarctic System in the United Nations has been that the continent is not a cornucopia of minerals, and that exploitation is far-off and may never be economical. By contrast, De Wit considers that the Dufek platinum mine may now be a more realistic venture than, for example, the north-west Pacific manganese nodule sites. His arguments require serious consideration by all concerned with the minerals regime negotiations.

Shapley's book is a view of Antarctic politics from the United States, with special emphasis on resources. Discussion of the interests of other countries is generally brief, but the examination of past and current American policy is detailed and useful despite the author's tendency to accept the views put forward by the United States government on matters such as the observance of the Treaty. The American approach in the past is criticized as ambivalent and *ad hoc*; indeed that nation's failure to grasp the nettle of deciding precisely what areas it was interested in not only carried through right up to the negotiation of the Treaty but still exists. Publicly available documentation shows this indecisiveness in the post-war period with wholesale claims on the basis of papers prepared after flights over the vast areas concerned.

The official view of the United States government, given by Shapley, is that the Antarctic Treaty and its related measures have been observed in letter and in spirit, as demonstrated by American inspections. This contention, which was strongly advanced in the defence of the Treaty System at the United Nations, has been challenged and requires examination. Among the recent examples of disputes are the French airstrip at Pointe Geologie, which has affected local fauna and has apparently even been the subject of protests by the Australian government, and

the environmental effects of the Chinese Great Wall Station. These instances of alleged breaches, taken with the past record of demonstrated specific cases, are of special interest because the Consultative Parties have placed heavy emphasis on their responsible custodianship of the Antarctic environment and have undertaken that the minerals regime will ensure that there are no serious effects on the environment.

The current Antarctic Treaty runs out in 1991, seen by Shapley as the hour of reckoning. However it is unlikely that the Treaty will be terminated then or that any of the Consultative Parties will seriously review their participation; even the most vocal claimants would have difficulty improving their position outside the Treaty. But 1991 will be a landmark because the Treaty then becomes open-ended. By that date an agreement over minerals will presumably have been negotiated, and the developing nations initiative in the General Assembly, aimed at a "common heritage" regime, will have been defeated. However the UN move has already had a substantial impact — for example in the granting of observer status to the ordinary parties to the Treaty who previously had no real rights — and Shapley presents a useful discussion of some of the issues at the UN.

The author's analysis of current American policy raises the fundamental question of what the United States is trying to do in Antarctica. Criticisms of the National Science Foundation's continuing emphasis on basic research should be taken up. But one must also ask why the United States should spend substantial funds on krill research when it is clear that American fishermen are unlikely to harvest the resource. Environmental protection would justify a modest programme, but large-scale research could be seen as a gift to Soviet krill exploitation.

Shapley points out the constraints of logistics in the American programme and it may be asked whether the huge cost of support activities is justified when compared with other nations' budgets.

She argues that the future of the Antarctic System may require a compromise with the developing nations and the UN. But the Consultative Parties include the superpowers, other major developed nations and some leading developing countries (India, Brazil, Argentina, China). The Consultative Parties have given no past indication of any real desire for compromise with the advocates of Antarctica as the common heritage of mankind, and it would be surprising if this notion would be accepted in the minerals negotiations. Shapley points out that the anticipation of viable mineral resources (as distinct from economic realities) is politically important and this has been demonstrated for the deep seabed. Current offshore research by the Consultative Parties, characterized by them as scientific research, is resource-orientated and is indeed playing with fire. A promising find might even wreck the minerals negotiations. □

Francis Auburn is an Associate Professor in the Law School, University of Western Australia, Nedlands, Western Australia 6009, Australia.

Institute of Oceanographic Sciences

NEW FROM

ADIU

Armament & Disarmament Information Unit

NATO CHEMICAL WEAPONS POLICY AND POSTURE

by Julian Perry Robinson

The fullest account yet published of the history and recent development of chemical weapons policy in NATO, by a leading specialist in the field. The paper documents the history of NATO CW doctrine, weapons and training, US Congressional actions on binary chemical weapons, and the adoption of a NATO Force Goal for binaries in May 1986. Detailed appendices set out the positions of NATO member governments and parliaments on the chemical weapons force goal.

ADIU Occasional Paper No. 4
 ISBN 0 950 177 00 0 (hbk) £15.00
 ISBN 0 950 177 01 7 (pbk) £7.50

ALSO AVAILABLE

ADIU REPORT

The bi-monthly journal of the Armament & Disarmament Information Unit. Unique coverage of arms control developments, British parliamentary proceedings and European defence and disarmament affairs, plus a comprehensive listing of recent publications in the field.

ISSN 0 950 177 02 4 (hbk) £15.00
 ISSN 0 950 177 03 1 (pbk) £7.50

A full list of ADIU publications is available on request.

Orders for publications should be sent, with remittance payable to
UNIVERSITY OF SUSSEX,
 to The Secretary, ADIU, SPRU, Mantell
 Building, University of Sussex, Brighton,
 East Sussex BN1 9RF, UK.

Building bridges between lives

Shirley Lindenbaum

Return to the High Valley: Coming Full Circle. By Kenneth E. Read. *University of California Press*: 1986. Pp. 269. \$18.95, £15.95.

KENNETH Read, the first anthropologist to work for an extended period in the New Guinea Highlands, has already provided one distinguished account of the quality of life among the Gahuku-Gama of the Asaro Valley. *The High Valley* (1965) took as its central themes the relations between men and women and the structuring of relationships between and within generations, although there was little evidence of an interest in anthropological theory as such. Instead, Read's search for the individuality of others and for shared moral concerns led to complex portraits of village leaders, male age mates, and the harsh and poignant ceremonies of initiation and marriage. The book was also a subjective record of self-discovery.

Return to the High Valley joins personal and professional past to present. Based on two short visits in 1981 and 1982, 30 years after the original research, it is a post-colonial work in a variety of ways. Read demonstrates his unique capacity to convey the momentous changes that have recently taken place in New Guinea, aware that those born after 1950 experienced the events in a different way from those men and women who grew up in more traditional times. The original themes thus surface again, played out now against the backdrop of the past. The transformation is captured in detail, vast and small: in the spread of a cash economy and new forms of employment, in the demise of the old ritual and ceremonial exchanges of wealth, in the growth of Goroka from a small government settlement to a town of over 10,000 people with a teachers' college and discotheque, and in the rumble of passing traffic now augmenting the small waking sounds of a rural community.

The book's post-colonial stance is also in the method of its telling. Aware of the unanticipated publicity showered on the subjects of the first volume and the ability of a younger generation to read past and present texts, the contours of the contemporary tribal political landscape are merely sketched, and detailed portraiture is reserved for persons who are dead, or for the members of Read's Gahuku family — the relatives of Makis. Makis, the most influential man of the tribe, and Read's companion, elder brother, friend and mentor from 1950 to 1952, died in a road accident in 1969. The "subject" of Chap-

ter 2 in each book, Makis is a more constant presence in the second work, due in part to changing conventions in ethnographic writing. Although *The High Valley* was in many ways a precursor of current trends which attempt to find new ways to represent the voice of informants adequately, *Return* replaces the subjective focus of the earlier autobiography with the view that Makis is the lens through which Read acquired some comprehension of the subtleties of Gahuku life. The life observed and written about is said to be in large part a distillation of Makis, leavened, of course, by the author's own perspectives and professional training. This more complex appraisal gives authorship not to the monologue of the observing anthropologist, but to the many unwitting local contributors to the text, and conveys more accurately the way anthropological knowledge is generated in intense, interpersonal engagement.

Critical at times of some Gahuku cultural practices and of the behaviour of certain individuals, and disaffected by the

embrace of what appear to be the most vulgar elements of our own society's offerings, *Return to the High Valley* is, nonetheless, a work of elegant cultural translation. It is the product of a mature craftsman who confesses a love for words, for the rhythm of sentences juxtaposed within a paragraph and a delight in visual properties. The special quality of the relationship between Makis and Read is part of the magic of the book. Although diffident about the ability of outsiders to penetrate the inner worlds of others, and aware of the impediments posed by attempting to convey such understandings in the language available to us, Read's account is a testament to the creation of kinship across the boundaries of different ways of life. □

Shirley Lindenbaum is Associate Professor of Anthropology, Graduate Faculty, New School for Social Research, 69 Fifth Avenue, New York, NY 10003, USA. She is also Editor of the *American Ethnologist*, and has made two field trips to the Eastern Highlands of Papua New Guinea.

Potential of plants

W.A. Hughes

Plant Cell Culture Technology. Edited by M.M. Yeoman. *Blackwell Scientific*: 1986. Pp. 375. \$52.79, £34.50.

THE evolution of techniques for plant-tissue culture into a viable technology has been slower than might have been expected. This is due, in part, to the inherent plasticity of plant cells that defies most attempts to lay down general rules of behaviour at either the cellular or tissue level. In *Plant Cell Culture Technology*, M.M. Yeoman has brought together a number of leading exponents to review selected aspects of this topic. The declared aim is to provide a readable, up-to-date account (at least to early 1985) of the commercial impact and potential of plant cell and tissue culture. Emphasis throughout the book is very much of the cellular rather than molecular aspects.

The book starts particularly well, with the first 100 pages devoted to vegetative propagation. These chapters look at the role of determination (Meins), a review of general methodology (Hussey) and finish with the implications for plant breeding with particular reference to somaclonal variation (Scowcroft and Ryan). The main body of the book is taken up with several chapters dealing with those techniques that have not yet been fully exploited. The uses of cryopreservation and gene banks (Withers) is described in much detail, as is the importance of cell selection (Dix) and the perpetual attempts to make developing secondary products a profitable pro-

cess (Fowler followed by Lindsey and Yeoman). Although there is much of interest here, these lengthy chapters (especially those by Withers and Dix) put the book out of balance. Indeed the contribution on genetic manipulation and cell transformation (Draper, Davey and Freeman) is relegated to the last 30 pages, and even then the focus is primarily on specific delivery systems for DNA into plant protoplasts. In an area where tissue culture may arguably yet play its most important role, I felt their view was unnecessarily limited.

Each of the authors provides a brief introduction to their area for non-specialist readers; for the researcher there is an excellent bibliography with titles. The tone throughout the book is mostly realistic — the veneer of commercial application is a thin one and few, if any, attempts are made to hide this fact. The mandatory introduction and outlook are written by the editor and clearly reflect the cautious optimism of the intervening chapters. He reminds the reader that it was eight years ago that the late H.E. Street looked to the future in an earlier Botanical Monograph (*Plant Tissue and Cell Culture*, 1977). The problems outlined then are still with us. However, the selective choice of topics in this book has not caught the present excitement in this area. Nevertheless, there are enough ideas here to act as a catalyst for the next eight years and as such could be read profitably by serious students, young and old alike. □

W.A. Hughes is a research scientist working on the biochemical aspects of plant cell differentiation at Unilever Colworth Laboratory, Sharnbrook, Bedford MK44 1LQ, UK.

Coping with the human factor

The Soviet authorities displayed considerable candour in Vienna last week. "Gross human error" was blamed for the Chernobyl accident.

Now the evidence presented at the meeting is being sifted for the lessons that will need to be well learned if the industry is to have a future.

Search for extra safety

By the time the operators of the Chernobyl Number 4 reactor pressed the SCRAM button to close it down it was already too late. That is one of the conclusions of the modelling studies by which Soviet scientists have simulated the course of events leading up to the accident. The simulations show that uranium oxide fuel elements in the upper part of the core had already begun to disintegrate as a consequence of the high temperatures they had reached. In the event, the downward movement of the control rods actuated by the SCRAM button, called AZ-5 on the control panel, was halted half-way along their track, presumably by a mechanical obstruction of some kind. Although the operators disconnected the control rods from their servo-motors with the objective of letting them fall freely, within 20 seconds an explosion within the reactor vault ended the useful life of the reactor.

By the account given by the Soviet delegation in Vienna last week, this was the culmination of 24 hours of frustration in the Chernobyl control room, whose occupants were 12 hours later than planned with a live test of a scheme for extracting emergency electric power from one of the station's two turbo-generating sets. The test had been carried out on two previous occasions, in 1982 and 1984. The plan, on 25 April this year, was to use the impending shut-down of the Number 4 reactor

(for routine maintenance) as an opportunity for a further test.

The objective was laudable enough, to bridge the gap that could arise if a reactor was accidentally isolated from the power grid. Even shut-down reactors need external power for circulating coolant that will carry away heat from radioactive fission products accumulated in the fuel. At Chernobyl as at other reactors, there were stand-by arrangements for supplying emergency electric power, storage batteries and three 5,500 kW diesel generators. But the storage batteries would have sufficed only to keep the instrumentation, control panels and some control systems operating, and would not have been powerful enough to drive the cooling pumps.

Using the energy of a decaying turbine set, which may amount to several megawatt-seconds, to fill this gap seems a prudent preparation for emergencies. The technical problem is to ensure that the power is supplied at a sufficient voltage (6 kV at Chernobyl) for as long as possible. According to Soviet statements in Vienna last week, there is, even so, "much discussion in the Soviet Union of the necessity of these tests".

The planning of the tests seems to have been in tune with the general sloppiness of the operation of the control room at the end of April. Because the two earlier tests had shown the voltage to fall more rapidly

than expected, the electromagnetic loading of the generators had been modified. But the test planned for 25 April was otherwise to be a straightforward repetition of the two earlier tests. The planning of the tests had been delegated to a group from an unidentified "electrotechnical institute" which is apparently not an academic institution but a group of engineers functioning as consulting engineers in the Western sense.

The specification of the test procedure was described as "poor". Specifically, the section of the planning document dealing with safety measures had been "drafted in a purely formal way". The planners had not thought through the dangers that might arise and specified remedies, but had instead advised that, in case of an emergency, the staff should act "in accordance with plant instructions".

In spite of its intrinsic defects, the plan for the test survived the scrutiny that, according to Academician V.A. Legasov, head of the Soviet delegation, is customary on these occasions because it bypassed the normal procedures. In particular, the plan was not shown to the design group responsible for RBMK reactors, which is a "violation of regulations". Because the plan required that the emergency core cooling system should be shut off, the implication is that it would have been given short shrift if it had gone beyond the chief engineer's desk.

The Soviet accident report says that because safety considerations had been skimmed in the planning of the test, "the staff involved were not adequately prepared . . . and were not aware of the possible dangers". This is not an absolution for the staff, which is charged not merely with departing from the programme but with breaking the station's safety regulations in the process.

In the sequel, the catastrophic explosion of the reactor, the test of the spinning-down of the turbo-generator as a source of emergency power featured in two ways that are not fully explained by last week's statements. First, the operating crew appears to have believed that it might be necessary to carry out at least one repetition of the test (which is why an attempt was made to keep the reactor operating at low power). Second, the crew seems not to have realized that they could have carried through a single test at any time, thereby avoiding trouble, simply by allowing the safety system to close down the reactor. □

Associated Press

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Chronology of a catastrophe

CONFUSION in the Chernobyl control room could only have been increased by the way in which the planned test was delayed. The first steps to close down the reactor for its planned maintenance were taken at 01:00 on 25 April, when the staff began to reduce the thermal power output of 3,200 MW. Exactly twelve hours later, power had been reduced by a half, and preparations were made for shutting down one of the reactor's two generators (Number 7 in Chernobyl notation) by connecting the station's internal equipment to the output of the other (generator 8). The emergency core cooling system was disconnected an hour later (apparently as required by the test programme). Then came the first hitch in the plan.

By the Soviet account, the Kiev regional controller of the Soviet electricity grid asked that the planned shutdown of the Number 4 reactor should be delayed for some hours to accommodate an unexpected demand for electric power. The station complied, and the reactor continued to operate at half-power and with its emergency core cooling system disconnected. In the event, the interruption lasted for just over 9 hours. Abagyau insisted last week that this prolonged period of operation at half-power did not contribute directly to the accident, but it may well have meant that a less knowledgeable control-room shift took over. Whatever the case, events began to go awry once the station was free to begin the test programme, at 23:10 on 25 April.

For an hour and 18 minutes, Soviet statements say, the operators continued to reduce the power output of the reactor, aiming at the range between 700 and 1,000 MW of thermal power specified in the test plan. At some point, the operators switched from the automatic system for regulating the reactor power by the independent regulation of the power density in each of twelve zones in the reactor core. This automatic system is designed to make the power distribution in the reactor uniform, which is beneficial both for the stability of the system and the economic use of the fuel. Regulation is by the automatic movement of neutron-absorbing control rods containing boron, of which there were 211. This system of local control is normally used only at high power, however. When a RBMK-1000 reactor is being started up, the control rods are driven (again automatically) by a system in which the power output of the reactor as a whole is the criterion, so that it is natural that the operators should have switched from the first to the second system of control (at 28 minutes past midnight on 26 April) when they were on the way down again.

This was the point at which the operators made the error from which all the

others flowed. The Soviet statements say that the operator neglected to reset the parameters of the automatic control system correctly. The consequence was that the target power level of 700–1,000 MW was vastly overshoot. The power output of the reactor fell below 30 MW, virtually on the edge of shut-down.

At this point, the test should have been abandoned. The Soviet statements say that the decision to carry on was not merely a violation of standard operating procedures but even of the written test-plan, inadequate though that may have been. The explanation of their perverse decision last April is that they were aware that, if the test could not be done then, they might have to wait another year, for another planned shut-down of their reactor, before they could tell whether their modified equipment could sustain the voltage of a decelerating turbo-generator set.

What followed was tragic-comedy. At this stage, the reactor would be badly

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Legasov points the finger

poisoned by xenon. Xenon-135 is a well-known reactor "poison" that, being a ready absorber of neutrons, can rob a reactor of reactivity, its capacity to sustain a chain reaction. In a reactor operating at a steady power level, there will be an equilibrium concentration of xenon determined by the balance between formation (through fission) and destruction (by neutron capture and radioactive decay with a half-life of 9.2 hours). While the physical equilibrium concentration of xenon is greater at greater reactor power, a rapid reduction of reactor power will, at the outset, saddle a reactor with an inappropriately large amount of xenon. The extreme case is when a reactor is shut down too quickly, when the poisoning may be so great that it cannot be started up again until the xenon has decayed naturally.

Nevertheless, the operators embarked on increasing the reactor power, for which purpose the automatic positioning of the control rods was overridden by manual control. In reality, to win more power it was necessary to withdraw most of the

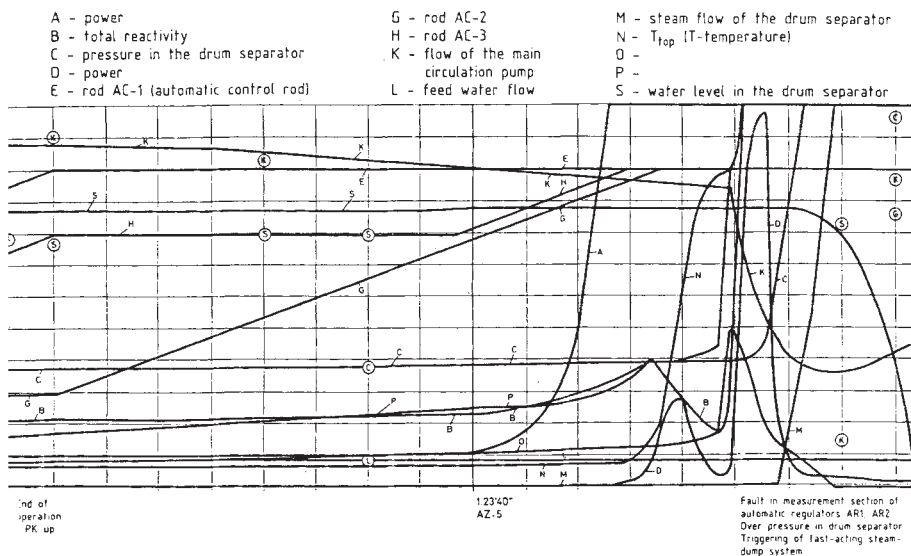
rods from the reactor to their full extent, another violation of station regulations.

The point was put graphically by Academician Legasov last week in Vienna, when he explained that the station regulations require that the operators should not, without reference to the chief engineer, operate the equipment when the potential reactivity of the reactor is less than the equivalent of thirty control rods. On this occasion, the chief engineer was not asked, even though the reserve of reactivity was far below the limit. The station regulations apparently also require that operators should never, in any circumstances, operate with less than the equivalent of 15 control rods of reactivity up their sleeves. "Nobody, not even the prime minister of the country" can give that permission. Yet, after the event, the reserve reactivity at Chernobyl can have been no more than the equivalent of 6–8 rods' worth of reactivity. To attempt the planned test in these conditions, according to Legasov, was like trying to conduct, in a high-flying aircraft, a test which involved opening the doors.

Somehow, by 01:00 on 26 April, the operators nevertheless managed to wring 200 MW of thermal power out of the reactor. Their decision to carry on with the test at a power so far outside the range of power specified in the lost programme was a further violation of the rules. But there is no plausible explanation of their subsequent decision to induce a massive increase in the flow of water through the fuel channels of the reactor by switching in the two stand-by pumps with which the reactor is provided, bringing the total to eight.

The Soviet report says that four of the pumps were to be driven by the spinning turbo-generator during the test, and that it was planned that four should remain in use after the completion of the first test. Others guess that the explanation is more complicated, and the operators, knowing that they were carrying out an invalid procedure, had miscalculated that more cooling water could only bring more safety.

Whatever the explanation, the consequences were immediate and portentous. Water was thundering through the reactor at a rate of nearly a ton a minute. The flow was so great that it might, in itself, have been sufficient to cause mechanical damage to the carefully engineered cooling channels, while at least some of the pumps were operating above their design capacity. Even by the time the first of the two stand-by pumps was switched into the cooling circuit, the temperature had begun to fall to the point at which steam was condensing into water. This had two effects — the balance between water and steam in the separator units was upset and the reactivity of the reactor was further decreased, entailing further withdrawal of the control rods to keep the power at 200 MW.



Monitoring shows the onset of problems. The "power" trace A is first to shoot off the scale.

On 01:07 on 26 April, the reactor was in an extremely sensitive condition, although the operators appear not to have appreciated the dangers. One of the Soviet delegation, the reactor engineer A.A. Abagyan, suggested last week that concentration on the electro-technical aspects of the planned test had made the operators indifferent to the nuclear aspects of their problem.

One result was that, when the water level in the steam separator drums began to fall, the operators disabled the devices designed to safeguard against a catastrophic loss of pressure in the water loops by shutting down the reactor. Instead, they used manual controls to top up the steam drums with cooling water. Within half a minute, by 01:19.30, however, the cold water reached the core, further reducing the temperature and the proportion of steam in the circulation, the automatic control rods reached their upper limits and the manual control rods were lifted out to their full extent. At this stage, the reactor was essentially uncontrollable, even if operating at low power and with a thunderous supply of cooling water.

With less than four minutes to go to catastrophe, the operating staff now shut off a set of steam relief valves on the cooling system, with the immediate objective of preventing the further decline of steam pressure, but with important consequences for the sequel which lay just a short while ahead. The course of events during these last few minutes has not been reconstructed from the station's records (which are said to be physically radioactive) but has been inferred from a computer model that has won general admiration for its sophistication and completeness.

With 1 minute 50 seconds to go, the operators appear to have realized that they would not restore the steam supply unless they moderated their pumping of cold water from the turbine condensers back into the main cooling circuit. The

calculation was correct. Within 30 seconds, steam began appearing in the system, a group of control rods moved inwards in response to the increased reactivity of the system and the pressure in the water circulation system began to rise after a further 20 seconds. On the face of things, the system must have seemed, at 01:22.30, to be on the way back to normal.

The truth was quite the opposite. The volume fraction of steam in the circulation was still increasing, adding to the reactivity of the reactor. By destroying the xenon poison, the increasing neutron flux worked in the same direction. Yet in this fool's paradise, the crew decided to begin their test. The operator in charge asked for a print-out of the reactor parameters, the

Crime and punishment

Q. What has happened to the operators?"

Prof. A.A. Abegyan: "They have been punished."

Q. "How?"

Abegyan: "That is not a field in which I am an expert."

steam supply to the spinning turbine was disconnected and the test begun at 01:23.0.

Possibly the operating staff knew by this time that their chance of restoring the reactor to normality, so as to have a chance of carrying out a second test, were slim. But if they were going to have to shut down the reactor soon, why not slip in the test while doing so? That is the most charitable expansion. But the operators had by then disabled the system that would have shut down the reactor automatically when it became apparent that neither of the turbo-generators was being supplied with steam.

The immediate consequence seems to have been a further increase of the circulation pressure, with the apparent paradoxical consequence that the reactivity of the reactor briefly declined as the fractional

volume shrank. The control rods, like yos, went up again, and then promptly down again. What should have given the show away (according to the model reconstruction) was the slow increase of the reactivity of the reactor above zero, which set in at 01:23.20 and which would have triggered shut-down if the safety system had not been disabled. In the seconds that followed, the steam fraction was also increasing, further adding to the reactivity and the neutron flux, but too quickly for the control rods of the automatic regulation system to be able to contain them.

The pressing of the panic button at 01:23.40 had no effect, in part because the control rods that were still available for use within the reactor were too few and too slowly moving. They were also too far out of the reactor core to have a quick effect. So, inevitably, there was a first surge of power 4 seconds after the button was pressed. It was now physically impossible for the cooling water to carry away the heat being generated in the fuel, with the result that the fuel began to disintegrate and the cooling water to boil. The consequence of that was the final and fatal surge of power, 5 seconds after the pressing of the panic button. Now, the cooling system had been rendered ineffectual by increased pressure whose effect was to close non-return valves in the circulation system. As the whole content of the circulation system turned to steam, something had to give. There is speculation that the graphite structure of the reactor core had already been distorted when the control rods were jammed mechanically within the reactor. Two seconds later, at what is estimated to have been 01:23.48, there came the first of two chemical explosions which lifted the roof of the reactor vault and threw red-lumps of graphite, together with pieces of uranium oxide fuel, onto the surrounding buildings and spaces.

There was much argument last week about the nature of the explosion, which had originally been described by Soviet sources as due to the explosion of hydrogen formed within the reactor. Now the opinion seems to be that the mere generation of steam at the great rate that characterized the last few seconds of the reactor may have been sufficient. The issue has a bearing on the question whether water-cooled reactors in general are potential sources of steam explosions.

For reactor engineers, Chernobyl must be a classic illustration for the textbooks, which at present go to great lengths to argue that the exponential growth that follows the appearance of excess reactivity in a reactor is potentially catastrophic. Now they will have a tangible and tragic example to quote. But the mystery remains that so many apparently well-trained people should hit on the precise blend of indiscretions that would make an apparently safe reactor blow up. □

Shutting the stable door

THE Soviet response to the accident at Chernobyl seems to have been both thorough and, after some initial hesitation, speedy. Participants at last week's meeting were almost fulsome in their expressions of admiration for an "impressive" feat of organization.

The 49,000 inhabitants of the town of Pripyat were given prophylactic potassium iodide tablets from 06:00 on the morning of 26 April and were told to stay indoors and not to drink fresh milk. This was done by members of the communist youth organization Komsomol.

Although the authorities in Moscow had been told of the accident when it happened, the clarity of the message was afterwards muffled by misleading signals from Chernobyl, suggesting that because the fires in the reactor building had been put out, the emergency was at an end. Even so, a high-level team of managers from the Soviet nuclear industry was assembled in Moscow and arrived in Chernobyl at 20:00 on the Saturday.

They must have been appalled at what they found. While the fires in the reactor building had been extinguished by firemen every bit as heroic as they are described in the Soviet press, the graphite core of the reactor was seen (by helicopter pilots) to be on fire; 10 per cent of the graphite, 250 tonnes in total, is estimated to have been consumed in this way. As a consequence of the high temperature of the fuel and the gas flow through the damaged core, the escape of radioactivity was virtually undiminished. And there were acute casualties to be dealt with.

The evacuation of Pripyat, 10 km from the site, was decided upon at 21:00 in the evening of the 26th, when the automatic monitoring stations showed that fallout had begun after a surprising delay (caused by the wind pattern), but not carried out immediately because of nightfall and the contamination of the prescribed evacuation routes. At this stage, the external intensity of radiation due to gamma radiation is estimated to have been between 600 and 1,000 R an hour. By the time that 1,000 buses had been assembled for the evacuation and allowing that the majority of the population will have spent much of their time indoors, it is estimated that doses of radiation due to gamma rays will lie between 1.5 and 5.0 rads, with a further 10–20 rads on account of exposure of the skin to beta-radiation.

With the continuing release of radioactivity, and the recognition that the rural population near the reactor would be less well-protected than those living in towns, the decision was taken ten days after the accident to evacuate all 135,000 people from a 30-km zone, removing the children to summer camps and sanatoria where

they could be given continuing medical supervision.

The treatment of those exposed to large doses of radiation in and around the reactor also won high praise last week, chiefly because of the accuracy with which the first damage assessments were made. The people exposed included firemen and reactor operators, including some who left the control room after the accident to see what had happened. One person was apparently killed within the reactor, where his body has not been found; another died of an excessive dose of radiation on the morning of 26 April at 06:00.

The treatment of the 350 people examined at hospitals within the first 36 hours was determined by the speed of onset of symptoms of radiation sickness (vomiting, headache and fever), and was afterwards confirmed by the associated fall of lymphocytes. People judged affected by radiation sickness were transferred either to specialized clinic Number 7 in Moscow or to a hospital in Kiev. The 20 classified as the most severely injured (of whom 17 died within 6 weeks) had received whole-body doses ranging upwards from 6 Gy (600 rem) to 16 Gy (1,600 rem).

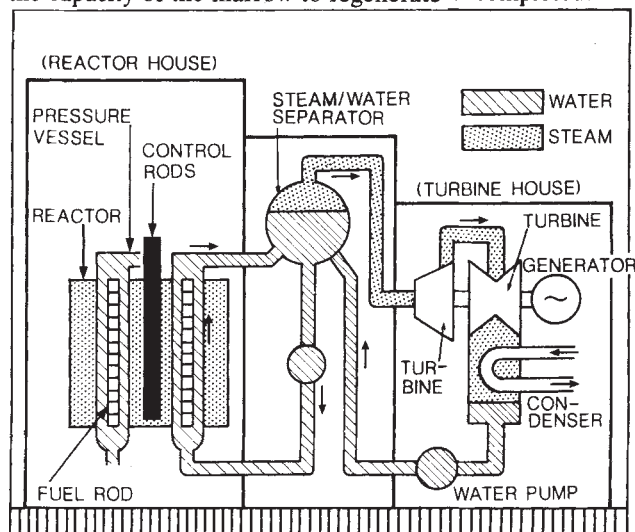
Two features of the treatment of those acutely exposed to radiation appear to have surprised Soviet physicians and, last week, those who heard their account of the past few months. First, a large proportion of those exposed to large doses of radiation were found also to have been damaged seriously by exposure of the skin to beta radiation, causing burns that were themselves hazards to life. Second, according to the Soviet report of the accident, bone marrow transplantation was of only limited usefulness because of the frequency of extensive burning in those most severely exposed while there is some evidence that in those less severely affected by radiation (with doses of about 600 rem) the capacity of the marrow to regenerate

spontaneously is sufficient to cause host-graft reactions (which may have contributed to the death of two patients). Platelet transfusion seems to have been invaluable in the prevention of bleeding in people with haematopoietic damage.

The treatment of the damaged reactor seems also to have been determined in the first few hours after the arrival of the team from Moscow during the evening of Saturday 26 April. Layers of boron carbide (to reduce the chance of a resurgence of nuclear fission), dolomite (as a heat sink and source of carbon dioxide) and metallic lead were dumped on the reactor from helicopters. With the continuing release of radioactivity, layers of sand and clay were added in the hope that they would absorb radioactive particles, but may in this respect have been counter-productive; their thermal insulation may have contributed to the increase of the temperature of the damaged core a week after the accident, and to the unexpected increase of the release of radioactivity then.

For the future, the recipe for the reactor is entombment. The final details have not yet been settled. For the present, vertical walls are being built around the reactor and those components in which fuel is thought to have lodged, the steam drums for example. The structure will be fitted with instruments for temperature and radiation measurements and provided with a closed-circuit gas-cooling system.

The future of the other reactors on the site remains uncertain, although the two operating reactors (called numbers 1 and 2) will probably be restarted when the problem of where the plant operators may live has been solved. The serious contamination of reactor Number 3, the twin of that damaged, appears to have been due to a delay in isolating the shared ventilation system; whether the reactor will be restarted remains to be decided, as does the question of whether the two reactors under construction at Chernobyl will be completed. □



The flow diagram for the Chernobyl reactor. A major deficiency in the design, Legassev acknowledges, is the tendency for the reactor's "reactivity" to increase when excess steam forms in the water-cooled fuel channels. The designs advantages are said to include simplicity and low cost, but a relatively complex system is required to keep the power output stable.

Tracking radiation release

THE wider radiological consequence of the Chernobyl accident may be less fearsome than was at first thought. This at least is the firm opinion of Professor D. Beninson, the Argentinian chairman of the International Commission on Radiological Protection who was the chairman of the appropriate technical session at the Vienna meeting.

The Soviet estimate that 50 MCi (say 2×10^{17} Bq) of radioactivity was released from the reactor in the form of non-gaseous fission products, with as much again in the form of noble gases (Ar and Xe) and iodine, which is based on measurements of fallout in the Soviet Union, has a formal uncertainty of 50 per cent. By this yardstick, the Chernobyl release was 100 times greater than that from the Windscale reactor fire in 1957 and more than 1,000 times as serious as the Three Mile Island accident two decades later.

So far, the Soviets have calculated only the external radiation doses likely to be incurred elsewhere in the Soviet Union on account of the passage of the radioactive cloud and the continuing effect of radiation from fallout lodged on the ground. The population evacuated from the neighbourhood of Chernobyl is estimated not to have been subjected to doses exceeding 2 rem with some exceptionally-placed individuals receiving between 30 and 35 rem.

For the 135,000 people in the evacuation zone, the collective dose calculated to have been received up to the point of evacuation is 1.6 million man-rem, which would imply an increase in the causation of cancers during the coming decades that would be less than 2 per cent or, using the ICRP estimate of 1.25×10^{-4} cases of cancer induction per rem of exposure, a total of fewer than 160 cases.

For the 75 million people in the western regions of the Soviet Union, the collective dose is estimated at 8.6 million man-rem up to the end of 1986 and 29 million man-rem over a period of 50 years. The Soviet report is quick to point out that the annual dose to the same population from natural radiation is 10 million man-rem, and that this is a field in which the assumption of a linear dose-response relationship may be least secure. With these reservations, the increased risk of cancer mortality in the western Soviet Union is estimated at less than 0.05 per cent.

The general opinion last week was that these estimates have been reached on conservative assumptions. (One interesting twist is that the frequency of spontaneous cancer, quoted as 10 per cent, is probably lower in the Soviet Union than in the West because of the lesser expectation of life in the Soviet Union.) Some evidence comes from whole-body scanning measurements carried out since the accident.

Estimates of the eventual dose arising from the intake of contaminated food are, by comparison, much less complete. On a calculation of the effect of caesium isotopes, of which 1 million Ci (of caesium-137) were released, the estimated dose-commitment over 70 years is estimated at

2.1×10^8 million man-rem for the Ukraine and Byelorussia alone, which would imply an increase of cancer mortality of 0.4 per cent. But there is some experimental evidence, based on whole-body scanning measurements, that the transfer of caesium through the food chain to the bodies of those exposed has not been as efficient as supposed, with measurements roughly 10 per cent of those predicted. □

What really went wrong?

Richard Wilson, author of an American Physical Society report on reactor safety, reflects on the lessons of Chernobyl.

LAST week's post-accident review meeting was historic for two reasons. It is the first time that there has been a nuclear power accident of the severity that we have all feared and it's also, perhaps, the first time since Ivan the Terrible that the Russians have discussed so openly the details of a domestic disaster. I believe that the second is the more important. Indeed, analysis of the event suggests that a primary cause of the accident was a lack of communication within the Soviet system. The meeting was therefore particularly difficult and important for the Soviets.

The Soviet delegation reporting on the Chernobyl catastrophe is led by Academician Legasov and includes Academician Ilyin, who directed the handling of the evacuation and subsequent study of radiological consequences; Dr Angelina Gusk'ova, who was in charge of most of the patients with radiation sickness; and many others directly responsible for various aspects of the disaster. This composition shows an unusual seriousness about reporting, helping and seeking help from the international technical community. The only notable exception was the absence of people who had operated RBMK-1000 reactors. The Soviets have worked fast since 26 April. Their report, hastily assembled though it is, is a monumental and impressive document. I, at least, would be proud to have been an author.

The Soviet RBMK-1000 reactors use a graphite moderator to slow down the neutrons, with uranium oxide fuel rods inside distinct vertical channels arranged in a lattice. Through the channels, coolant water flows and boils. There are two features to this reactor that are significant for major accidents. The first is that the moderator is solid and separate from the coolant. Removal of the liquid coolant, due to boiling, will not shut down the reactor. If other engineered design features fail to shut the reactor before an impending accident, the graphite and uranium must mechanically separate — and the ejection of an appreciable fraction of the fuel seems to be what terminated the nuclear

reaction at Chernobyl 4.

The other feature is that, in this reactor, removal of water in the channel (perhaps by an increase in steam content) leads to an increase of reactivity, a power increase, further boiling and removal (voiding) of water. Technically the reactor is then said to have a positive void coefficient of reactivity.

Of course this tendency to instability can be stopped by an engineered safety device — introducing boron absorber (control) rods to reduce reactivity faster than channel voiding increases it. Physicists prefer inherently safe devices, but admit the possibility of safe engineered devices, if proper attention is paid to them. This proper attention was not paid at Chernobyl and, in fact, the RBMK reactors have the worst shut-down capability of any commercial reactor in the world.

The control rods move quite slowly and, if in the "out" position, cannot reach the active zone, and so affect reactivity, for a few seconds. The designers therefore established a rule: at least 30 rods must be in the core at any time. Because view of the crucial importance of an engineered shut-down (scram) in this reactor, leaving this safety precaution to a rule, rather than a mechanical limit, or interlock, seems unwise.

At the power plant, an experiment had been planned to tell for how long a generator could continue to generate enough electricity to operate important systems after the steam to the turbine is cut off. There were delays; in order to allow the experiment to go ahead, the operators deliberately violated six safety rules. Two of these made the reactor exceptionally difficult to control and intensified the effect of the positive void coefficient. The coolant water was near saturation, so that a small change in hydraulic conditions could rapidly change the void fraction and thus the reactivity. These put the reactor time constant as low as a second, falling rapidly as a prompt critical situation was reached.

The reactor is not designed to control such small time constants. In three different ways, the operators disconnected the automatic reactor shut-down devices so that they could continue to operate in this unsafe condition. This meant that the

reactor could be stopped only by manual shut-down with all the delay involved.

The sixth error — disconnection of the emergency core cooling system — did not have any effect on accident initiation but the emergency cooling could have mitigated the consequences.

As a result of these errors, when the experiment began at 1:23:00 on 26 April, the reactor was uncontrollable and began a rapid power excursion. At 1:23:30, the power began to rise above the 200 MW operating level. At 1:23:50, the power had passed 320 MW and the period for power doubling was a second or so. At this point, the operators pressed the button AZ-5 to scram (shut down) the reactor. But it was too late. Before the control rods could reach the active zone, the power had gone over the 3,000 MW design power; by 1:23:48, water flow had stopped as the check valve closed; but, by 1:23:49, flow had restarted, presumably because the channels had ruptured. It appears that the control rods stopped before going completely in.

Both halves of the core had the same problems. It is therefore probable that most of the 1,640 fuel channels ruptured within a few seconds — and the reactor can stand the pressure increase of only a few channels rupturing at a time. Inevitably the ejected pieces of the core went through the roof, including both pieces of graphite moderator and pieces of uranium fuel, although whether there was an explosion in the technical sense of the existence of a shock wave is debatable. It should be noted that this is one location, and direction, where there is no containment and ejection of fuel elements like missiles is inevitable.

Why were the operators unaware of the danger? One can only speculate. The idea that one must have several rods partially in the reactor already to achieve a rapid shut-down is not immediately obvious and, to some people, including me, is counter-intuitive; one might think that the more rods that are out, ready to go in, the better. But the problem is the rate of insertion. Be this as it may, it is clear that the operators did not understand the reasons for the rules. One most important procedural requirement was also violated and had probably been violated before. Any experiment is supposed to be checked in detail with the director or the chief engineer of the station; but these staff considered the experiment to be merely an electrotechnical exercise, forgetting that all procedures in a reactor are interconnected.

I suggest that these errors are much more likely to happen in a compartmentalized society, as they have in the Soviet Union, than in the open society of the West. In the United States, a violation of these rules would surely be leaked to the *Washington Post* within a day or two, and

Reactor improvements

THERE is too much at stake for the Soviet Union to consider seriously abandoning its RMBK reactors. Quite apart from past investments in them, the need for power in the populous western regions of the Soviet Union could not otherwise be satisfied. Energy transport already accounts for 40 per cent of rail traffic in the Soviet Union, according to chief delegate Legasov. But the following improvements are to be adopted by 1987.

- The control rods will be fitted with new stops preventing the last 1.2 m of their present travel.
- Partly in compensation but also to give the reactor type a greater reserve of reactivity, the enrichment of the fuel will be increased from 2.0 to 2.4 per cent.
- Consideration is also being given to schemes for flooding the reactors with neutron-absorbing liquids or gases for rapid shut down in emergency.
- More instruments will be installed in the reactors. □

the perpetrators brought to account. For this reason, we must especially welcome the Soviet openness at the IAEA meeting. Let us hope it continues, not only between Soviet and US reactor designers but between designers and operators in the Soviet Union itself.

Detailed analysis of accidents, including human error, is common in the United States, particularly since Three Mile Island. The procedure is called "probabilistic risk analysis". It is vital that the Soviets use this procedure to the full to ensure that there are no other problems with these reactors that they have not taken into account. Dr Legasov said that "the fact that we have started later than many specialists in thinking about such matters is a fault on our part". Of course the Soviets are taking strong disciplinary and procedural steps to ensure that such a litany of human errors never again occurs. But in addition, several steps have been taken to make the equipment less liable to errors — even premediated ones.

As an interim measure, all RMBK reactors are being modified with limit switches so that the control and shut-down rods cannot be brought to a region where it takes several seconds for them to be effective. The rule that 30 rods must be partially in the core at all times is being modified to demand that 60 be in at all times. This not only makes shut-down faster but reduces the void coefficient. These changes are being made now to RMBK reactors and half are at present shut down for this purpose.

In a year or two, the fuel, which is now enriched to 2 per cent in uranium-235, will be replaced by fuel with 2.4 per cent enrichment. This will further reduce the void coefficient and enhance stability. In the

longer term, the Soviets are thinking about faster methods of shut-down such as gas or liquid injection of poisons.

The Canadians, who have experience with positive void coefficients, have offered in the design of shut-down systems with proper interlocks and key boxes to discourage unauthorized violations of regulations as much as possible. These should include simple prominent displays of computed excess reactivity and automatic shut-down if excess reactivity is too small.

There are no direct parallels between the RMBK reactors and light-water moderated reactors (LWRs). LWRs tend to shut down naturally, by boiling of the moderator, and although the Anticipated Transient Without Scram (ATWS) is serious for a boiling water reactor (BWR), there are still some hours for recovery rather than the 10 seconds in the RMBK. Even if six deliberate successive operator errors happen in the West, the hardware would make that less important. I therefore believe that the light-water reactors in the United States are safer than the RMBK 1000 reactors, even after the latter have been improved.

But immediate replacement of the RMBK reactors seems impossible — except perhaps by a return to coal-fired power plants. We must remember that the London air pollution incident of December 1952 caused 4,000 deaths in the first two weeks — which is much more than the 31 deaths and the 4,000 "projected" cancer deaths from Chernobyl — and it has been suggested that air pollution still causes thousands of premature deaths a year. So a simple comparison of risks is likely to suggest to a Soviet decision-maker that the RMBK reactors should continue to operate.

As General Secretary Gorbachev himself has noted, an accident in the Soviet Union affects the rest of the world. The Soviets have asked for help in understanding what steps to take and implicitly ask approval of steps already proposed. While these steps are excellent and deserve support, it was clear that the specialists at Vienna are reserving judgement on the safety of the RMBKs until it is clear that the Soviets are prepared to accept the detailed international scrutiny, both formal and informal, that occurs in the West. This must include exchanges of information and personnel including visits to the Soviet power plants. The compartmentalization of Soviet society must not be allowed to prevent developments such as these, and the outlook is promising. If we do not learn from our mistakes, we shall be condemned to repeat them.

Richard Wilson is Professor of Physics at Harvard University. The American Physical Society report on reactor safety was published in Reviews of Modern Physics earlier this year.

Detection of two satellites in the Cassini division of Saturn's rings

Essam A. Marouf & G. Leonard Tyler

Center for Radar Astronomy, Stanford University, Stanford, California 94305-4055, USA

The presence of two ~10-km-radius satellites within a 100-km-wide gap in the Cassini division of Saturn's rings is inferred from the observation of perturbations of nearby ring material. The perturbations are detected as regular fluctuations in the optical depth profile measured during Voyager radio occultation by the rings and are used to estimate the satellites' orbits and masses. The two satellites appear to 'shepherd' a 30-km-wide ringlet within the gap.

VOYAGER radio occultation by Saturn's rings has revealed complex structure in the Cassini division, comprising several distinct gaps alternating with regions of optical depth τ (at 3.6 cm wavelength) ≈ 0.2 – 0.4 . Lissauer *et al.*¹ suggested the presence of unseen embedded satellites whose gravitational torques are responsible for opening the division's gaps, but no such satellites were detected² in high-resolution images obtained by Voyager 2 of the two widest gaps.

We report here the detection of two embedded satellites believed to be orbiting within a Cassini division gap, ~100 km wide, located immediately outside the Huygens gap. We have observed wave-like optical depth fluctuations, which we interpret as satellite 'wakes' induced in streamlines of ring particles by systematic gravitational perturbations. The perturbations modulate the distribution of ring particles in the ring plane, and consequently the optical depth profile of the rings. The recent development of techniques to remove diffraction effects from the measurements permits the reconstruction of the optical depth profile with sub-kilometre radial resolution³, revealing the wake patterns of the two satellites.

Although at least 17 satellites are known outside the classical ring system, only one, located within the Encke gap^{3–6} in the outer part of Ring A, has been found previously within the main body of the rings. The presence of the Encke satellite was first recognized by Cuzzi and Scargle⁴ from observations of undulations in the Encke gap edge. This inference was strengthened by observations of wake patterns in Voyager stellar and radio occultation data^{3,5,6}, all of which yielded consistent results for the mass and orbit of the embedded object. The method used here to estimate the orbit and mass of the Cassini division satellites is similar to that used to study the Encke satellite wake.

Satellite wake model

The theoretical model used is based on the non-colliding streamlines approximation to perturbed orbits of ring particles, as developed by Borderies *et al.*⁷, whose notation we follow closely. To first order, the results are equivalent to those of the wake model developed from first principles by Showalter *et al.*⁵.

The orbit of the embedded satellite is assumed to be circular, with radius a_s . In a saturnocentric cylindrical coordinate system (r, ϕ) which rotates with the angular velocity Ω_s of the satellite, the perturbed orbit of a ring particle has the form

$$r = a[1 - e(a) \cos E(a)] \quad (1)$$

where a is the semi-major axis, e is the eccentricity ($e \ll 1$) and E is the eccentric anomaly. The values of e and E are given by^{7,8}

$$E(a) = \frac{2a[\phi + \Delta(a)]}{3|a_s - a|} = \frac{2a\phi}{3|a_s - a|} + \frac{\pi}{2} \quad (2)$$

and

$$ae(a) = 2.24 \frac{M_s}{M_p} \left(\frac{a_s}{a_s - a} \right)^2 \quad (3)$$

where

$$\phi(t) = \theta(t) - \theta_s(t_{\text{ref}}) - (t - t_{\text{ref}})\Omega_s \quad (4)$$

is the longitude in the rotating frame relative to the satellite, $\theta(t)$ is the longitude in the inertial frame, $\theta_s(t_{\text{ref}})$ is the longitude of the satellite in the inertial frame at some convenient epoch t_{ref} , t denotes time and M_s and M_p are the masses of the satellite and planet, respectively.

The variation of the phase angle $\Delta(a)$ in equation (2) with semi-major axis a is calculated using the impulse approximation for the satellite encounter. All particle orbits are assumed to approach the satellite along unperturbed circular paths and at $\phi = 0$ change in concert to the perturbed orbits described by equation (1). This is an idealization which does not apply in the neighbourhood of $\phi = 0$, but which is assumed to have little effect on the streamlines downstream from encounter (compare with Fig. 3a of ref. 9, calculated using numerical integration of an initially circular orbit). The loss of memory regarding previous encounters is assumed to be caused by collisions of streamlines and subsequent loss of eccentricity before a new encounter.

Downstream from encounter, streamlines interior (exterior) to the satellite orbit overtake (fall behind) the satellite, while differential orbit perturbations cause a region of previously uniform optical depth τ_0 to be coherently perturbed. Assuming $|a_s - a|/a \ll 1$ and $3|a_s - a|/|\phi a| \ll 1$, one can show that

$$\tau(a) \approx \tau_0 / |\partial r / \partial a| \approx \frac{\tau_0}{1 + \left(\frac{x_{\text{coll}}}{a_s - a} \right)^4 \cos \left(\frac{2a_s \phi}{3|a_s - a|} \right)} \quad (5)$$

where

$$x_{\text{coll}}^4 \equiv 1.49 \frac{M_s}{M_p} |\phi| a_s^4 \quad (6)$$

Note that when the change in ϕ during the period of observation is small, then x_{coll} becomes a constant parameter that characterizes the distance $|a_s - a|$ at which streamlines intersect. Thus, equation (5) applies only over the range $x_{\text{coll}} < |a_s - a|$. We refer to x_{coll} as the collision distance.

Salient characteristics of wakes as described by the model equation (5) are: (1) strong dependence of the optical depth maxima envelope on $|a_s - a|$, decreasing as $|a_s - a|^{-4}$ (the unbounded increase in $\tau(a)$ as $|a_s - a| \rightarrow x_{\text{coll}}$ is, of course, an artefact of the neglect of inter-particle interactions); (2) an optical depth minima envelope that remains confined to the range $\tau_0/2$ to τ_0 ; and (3) a quadratic wavelength variation with $|a_s - a|$ of the form

$$\lambda(|a_s - a|) = \frac{3\pi|a_s - a|^2}{a_s|\phi|} \quad (7)$$

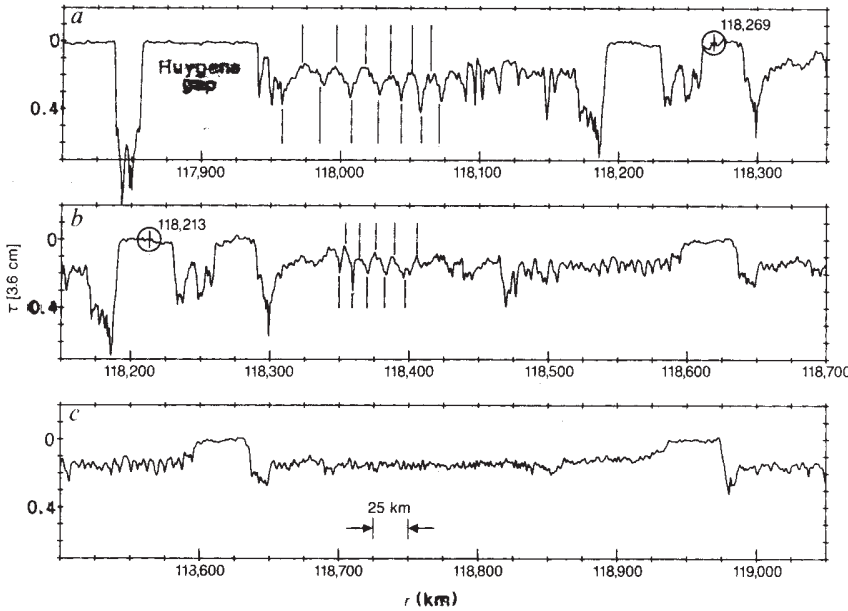
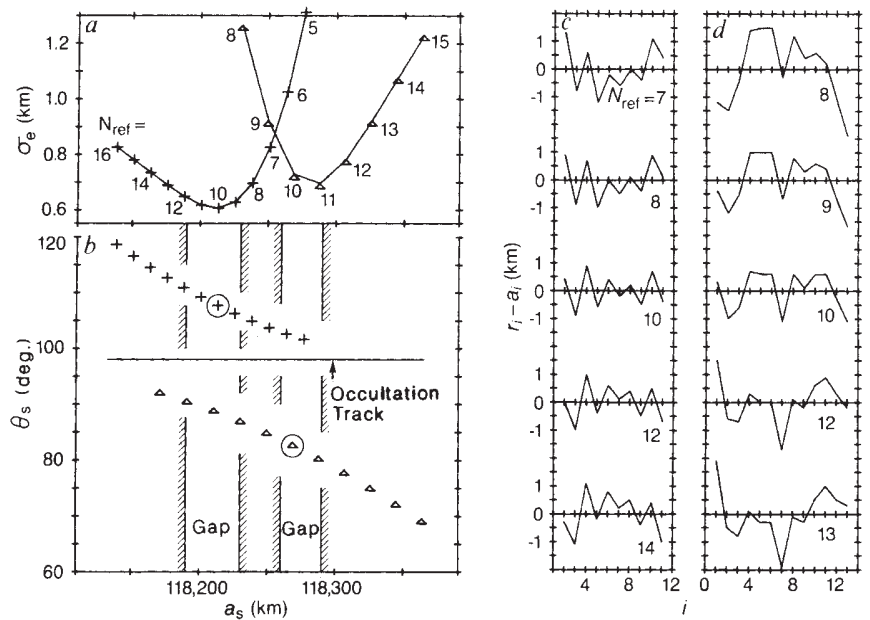


Fig. 1 Optical depth profile of the inner region of the Cassini division, reconstructed at radial resolution ~ 900 m from Voyager 1 radio occultation measurements at 3.6 cm wavelength. Panels *a–c* form one contiguous (overlapping) segment which includes the three 'plateaus' immediately exterior to the Huygens gap. The regular fluctuations at $\sim 117,950$ – $118,075$ and $\sim 118,350$ – $118,410$ km are believed to be the wakes of two satellites within the gap centred on $\sim 118,240$ km. In *a*, the circle within the gap identifies the orbital radius of the satellite responsible for the inner wake and the short vertical segments identify the extrema locations of the corresponding wake model. Similar notation is used in *b* for the outer wake. The two satellites may control the ~ 30 -km-wide ringlet within the gap. Panel *c* serves as a noise reference for *a* and *b*.

Fig. 2 Results of least-squares solutions for satellite orbits. *a*, Variation of r.m.s. fit error σ_e with orbit radius; discrete points identify a number of possible solutions, indexed by a reference integer N_{ref} : +, outer wake; Δ , inner wake. *b*, Corresponding longitude θ_s of the satellites plotted against orbit radius, a_s . The epoch assumed is $t_{\text{ref}} = 1980:318:04:54:39.39$ (UTC/ERT). The radial location of the gap and its embedded ringlet are identified by the vertical hatched lines. The two circled points on either side of the embedded ringlet correspond to the two most probable satellite locations. *c*, *d*, Illustration of the behaviour of the fit residuals of the outer (*c*) and inner (*d*) wakes for reference indices N_{ref} equal to, less than and greater than the most probable values.



when ϕ is nearly constant. This variation contrasts with that of linear density and bending waves, where the wavelength decreases in inverse proportion to distance from resonance¹⁰.

From equation (5), optical depth extrema occur at values a_i which are nearly independent of the satellite mass and are given by

$$\frac{2a_s|\phi_i|}{3|a_s - a_i|} = \pi(2N_{\text{ref}} + i), \quad i = 1, 2, 3, \dots \quad (8)$$

where N_{ref} is a reference integer introduced to identify an observed train of oscillations whose starting index is unknown, and odd and even indices i identify maxima and minima, respectively. In the observations discussed below, ϕ_i remains nearly constant over the duration of an observation. Assuming $\phi_i \approx \text{constant} = \phi$ in equation (8) and using $a_i = r_i$ at the locations of optical depth extrema (equation (1)), the observed values r_i can be used in equation (8) to constrain a_s , ϕ and N_{ref} . Equations (5) and (6) and the ratio of optical depths measured at successive pairs of extrema can then be used to constrain the mass ratio M_s/M_p . Solutions are ambiguous because the value of N_{ref} is uncertain.

Orbit and mass estimates

Figure 1 shows the profile at 3.6 cm wavelength of the inner Cassini division from Voyager 1 radio occultation measurements at ~ 900 -m effective radial resolution³. Figure 1*a–c* form one contiguous segment which includes four gaps, two of which contain embedded ringlets, bordering three 'plateaus'. The plateaus have comparable average background optical depth (~ 0.15 – 0.2) and comparable width (~ 250 – 300 km). The gap found to contain the two satellites is centred on radius $\sim 118,240$ km, shown in both Fig. 1*a* and *b*. The satellites may control the ~ 30 -km-wide ringlet located within this gap. The outermost plateau, shown in Fig. 1*c*, is nearly homogeneous and serves as a convenient indicator of the typical ring noise for the other two.

We argue below that the regular fluctuations apparent over the radial interval $\sim 117,950$ – $118,075$ km are actually the leading wake of a satellite located at radius $\sim 118,269$ km. Similarly, we argue that the regular fluctuations apparent in Fig. 1*b* over the interval $\sim 118,350$ – $118,410$ km are the trailing wake of a second satellite located at radius $\sim 118,213$ km. Other regular, but more subtle, fluctuations are evident interior to the narrow gap at $\sim 118,615$ km; their cause has yet to be identified.

Table 1 lists relevant measurements used in our model fit. Ideally, one would like to use the full observed pattern to fit the model. Unfortunately, this approach is hampered by variations in the background opacity across the observations and by significant deviations in the detailed shape of many individual observed oscillations from the wake model, equation (5). Consequently, only the estimated amplitudes τ_i and locations r_i of optical depth extrema are used. Almost all τ maxima observed are well localized; however, several of the observed minima are relatively broad and require interpolation to estimate their locations. The exact error in individual estimates r_i is not well determined but is bounded by a worst-case value of approximately ± 1 km.

Table 1 indicates that observation of the inner wake lasts only ~ 1.45 s, during which time the longitude θ_i changes by $\sim 0.033^\circ$.

Table 1 Relevant measurements and model fit results

Wake	i	t_i^*	Measurements			Model fit results	
			$\tau_i^\dagger$	$\theta_i - 98^\circ$ (deg.) [‡]	r_i (km) [§]	a_i (km)	$\frac{M_s}{M_p} \times 10^{12}$ [¶]
Inner	1	0.821	0.45 ± 0.06	0.1345	117,957.5	117,957.8	—
	2	1.017	0.13 ± 0.01	0.1298	117,973.0	117,972.0	31.8 ± 4.2
	3	1.175	0.27 ± 0.02	0.1261	117,985.5	117,984.9	23.0 ± 3.7
	4	1.307	0.15 ± 0.01	0.1229	117,996.0	117,996.7	24.3 ± 2.5
	5	1.446	0.33 ± 0.02	0.1196	118,007.0	118,007.6	22.0 ± 2.3
	6	1.573	0.15 ± 0.01	0.1166	118,017.0	118,017.6	17.2 ± 1.8
	7	1.712	0.32 ± 0.02	0.1133	118,028.0	118,026.9	9.9 ± 1.7
	8	1.800	0.20 ± 0.01	0.1112	118,035.0	118,035.6	9.9 ± 1.6
	9	1.907	0.35 ± 0.03	0.1086	118,043.5	118,043.6	10.5 ± 1.5
	10	1.996	0.18 ± 0.01	0.1065	118,050.5	118,051.1	10.9 ± 1.0
	11	2.084	0.42 ± 0.03	0.1044	118,057.5	118,058.1	10.7 ± 1.0
	12	2.179	0.17 ± 0.01	0.1022	118,065.0	118,064.7	7.4 ± 0.8
	13	2.267	0.35 ± 0.02	0.1001	118,072.0	118,070.9	—
Outer	11	5.777	0.19 ± 0.01	0.0168	118,350.0	118,349.6	4.0 ± 0.5
	10	5.821	0.06 ± 0.01	0.0157	118,353.5	118,354.2	6.2 ± 0.4
	9	5.897	0.30 ± 0.01	0.0139	118,359.5	118,359.0	4.2 ± 0.4
	8	5.953	0.12 ± 0.01	0.0126	118,364.0	118,364.2	3.0 ± 0.5
	7	6.029	0.20 ± 0.01	0.0108	118,370.0	118,369.8	3.8 ± 0.6
	6	6.099	0.11 ± 0.01	0.0092	118,375.5	118,375.9	5.0 ± 0.8
	5	6.193	0.21 ± 0.01	0.0069	118,383.0	118,382.4	6.0 ± 0.8
	4	6.263	0.105 ± 0.01	0.0053	118,388.5	118,389.4	7.5 ± 1.1
	3	6.382	0.21 ± 0.01	0.0024	118,398.0	118,397.1	10.0 ± 0.9
	2	6.471	0.09 ± 0.01	0.0004	118,405.0	118,405.4	—

* Time in seconds relative to 1980:318:04:54:35.0 UTC/ERT.

† Optical depth extrema; odd and even indices i identify maxima and minima, respectively.

‡ Longitude in inertial frame measured anticlockwise from the intersection of Earth's mean Equator (1950.0) with the ring plane at the ascending node (axis Ω in Fig. 5).

§ Estimated radial locations of optical depth extrema. The estimates are subject to a worst-case error of approximately ± 1 km. Data of resolution varying from 300 to 900 m were used to obtain these estimates, depending on the signal-to-noise ratio of the feature. Absolute radial distances are subject to an uncertainty of about ± 10 km.

|| Theoretical radial locations of optical depth extrema for the most probable satellite location, schematically shown in Fig. 5.

¶ Satellite mass (relative to Saturn) estimated from fitting ratios of successive values of optical depth extrema.

The longitude of a satellite somewhere within, or close to, the gap at 118,240 km would change by only 0.013° during the same period. Thus, the relative longitude ϕ as given by equation (4) is approximately constant during this period. The same is also true during the ~ 0.7 s observation time of the outer wake. For either observed wake and a given N_{ref} , the satellite orbit radius a_s and longitude θ_s at some epoch t_{ref} is estimated from a least-squares fit of the theoretical predictions of equation (8) to the observed extrema locations r_i . We refer below to the standard deviation of the residuals $r_i - a_i$ as the r.m.s. error σ_e .

Figure 2 summarizes the results of the fit. For each assumed

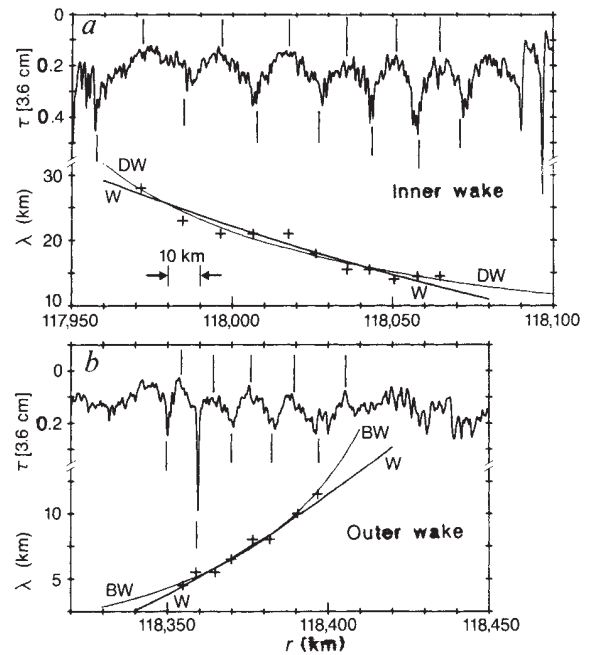


Fig. 3 *a*, Observed inner wake at 400-m resolution; the short vertical segments identify the theoretical extrema locations of the most probable wake model. Below, the theoretical wavelength profile of the wake (*W*) is superimposed on the measurements (+) to illustrate possible systematic deviations. The other solid curve (*DW*) is obtained assuming the observed fluctuations to be a density wave. *b*, As *a*, but for outer wake; *BW*, bending wave.

N_{ref} , the estimates a_s and θ_s and the corresponding error σ_e are grouped into two pairs (a_s, σ_e) and (a_s, θ_s) and plotted as the discrete points shown in Fig. 2*a* and *b*, respectively. Individual points are labelled by the corresponding N_{ref} in Fig. 2*a*, and the symbols '+' and 'Δ' are used to denote results for the outer and inner wake, respectively. Hatched vertical lines are drawn in Fig. 2*b* at the radial locations of the edges of the gap and the embedded ringlet.

Consider first the outer-wake results (+). The error σ_e in Fig. 2*a* exhibits a concave behaviour, with the minimum falling within the inner region of the gap. The minimum is rather broad, however—a consequence of the finite observed width of the wake. Unfortunately, an objective test of the statistical nature of the residuals, to determine which of the solutions obtained, if any, are acceptable fits, is not possible, due to the small number of extrema observed. We therefore approach this problem heuristically, by inspecting the behaviour of the residuals as N_{ref} varies over the range of values indicated in Fig. 2*a*.

Figure 2*c* shows the residuals for outer-wake model fits plotted against the extrema index i for five values of N_{ref} centred on the most likely value, $N_{\text{ref}} = 10$ (see Fig. 2*a*), as labelled. The clear curvature exhibited by the residuals corresponding to $N_{\text{ref}} = 7$ and 14 in Fig. 2*c* is indicative of strong correlation among adjacent indices, and hence unacceptable fits. The curvature is smaller for $N_{\text{ref}} = 8$ and 12 but is nonetheless still evident. For this wake, we therefore take the range $9 \leq N_{\text{ref}} \leq 11$ to be the range of acceptable models under the measurement constraints. Within this range, the model $N_{\text{ref}} = 10$ defines the most likely satellite location to be $a_s = 118,213$ km and $\theta_s = 107.7^\circ$, the circled '+' within the inner gap in Fig. 2*b*. The other two models, $N_{\text{ref}} = 9$ and 11, define an ambiguity in a_s and θ_s of approximately ± 12.6 km and $\pm 1.6^\circ$, respectively. This ambiguity is much larger than the formal statistical errors in a_s and θ_s of approximately ± 2.9 km and $\pm 0.17^\circ$, respectively, calculated assuming ± 1 km uncertainty in the observations, r_i .

A similar procedure is followed for the inner wake. The minimum error σ_e falls within the outer region of the gap, as

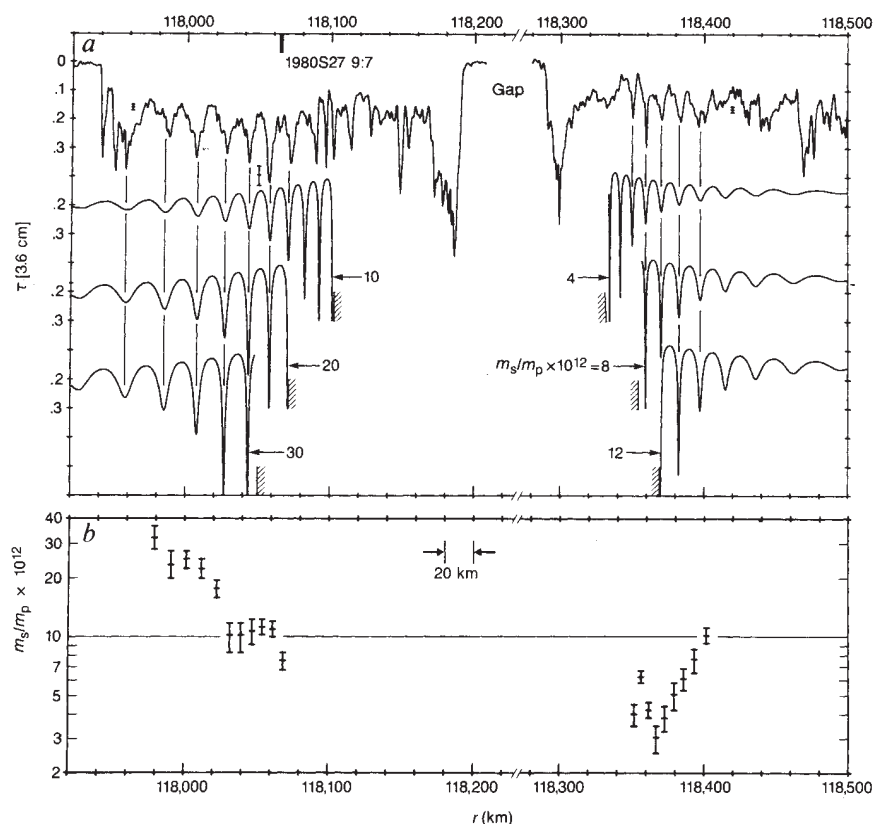


Fig. 4 *a*, Measured opacity profiles of the inner and outer wakes (top), compared with theoretical wake profiles, each computed assuming the most probable satellite location and the mass ratio labelled. For each wake, three theoretical profiles represent approximate lower and upper bounds on the mass, as well as an intermediate case. The short hatched vertical segments identify the onset of crossing streamlines. No single theoretical profile was found to fit simultaneously the amplitudes of all oscillations observed. *b*, Mass inferred from matching individual observed oscillations. The variability of mass across the wake is possible due to the neglect of collective effects in the model.

evident in Fig. 2*a*, where acceptable fits are limited to the two models $N_{\text{ref}} = 10$ and 11. Of the two, we take the first to define the most likely satellite location $a_s \approx 118,269$ km and $\theta_s \approx 83.2^\circ$, partly on the basis of its location within the gap, the circled ' Δ ' in Fig. 2*b*. The ambiguity in a_s and θ_s in this case is approximately 19.1 km and -2.3° , respectively, again values significantly larger than formal statistical errors of approximately ± 2 km and $\pm 0.12^\circ$, respectively. In the case of the inner wake, however, some degree of correlation between adjacent residuals persists even for the most likely model, $N_{\text{ref}} = 10$, as can be seen in Fig. 2*d*, indicating possible systematic deviation from the assumed theoretical model. Further evidence for such deviation is apparent in the wavelength profile discussed below.

Figure 3*a* and *b* show a more detailed view of the two wake profiles at 400-m resolution. As in Fig. 1, the short vertical line segments identify the extrema locations a_i corresponding to the estimates a_s and θ_s obtained; numerical values of a_i are given in Table 1. Also plotted in Fig. 3 are discrete points showing observed wavelengths versus distance, where wavelength is computed from Table 1 as $(r_{i+1} - r_i)$ and plotted at the mean radius $(r_{i+1} + r_i)/2$. Superimposed on the discrete points is a theoretical solid curve labelled 'W' (for wake), computed using equation (7) and assuming the most likely estimates a_s and θ_s obtained. Although general agreement with the data points is evident, the discrete points in Fig. 3*a* appear to 'weave' around the theoretical curve in a systematic manner. The cause of this behaviour, if significant, is unknown.

We have also considered linear spiral density and bending waves as the possible source of the observed optical depth fluctuations. Least-squares fits of such phenomena to the wavelength data using standard techniques^{11,12} lead to the two curves labelled 'DW' (for density wave) and 'BW' (for bending wave) in Fig. 3. In both cases, it is clear that, over the limited range of wavelength variation observed, the data are insufficient to discriminate among wake and wave patterns. The density-wave theory leads to a resonance at a radius of $\sim 117,879$ km (within the outer region of the Huygens gap in Fig. 1) and implies a surface mass density of $57.5 (m-1) \text{ g cm}^{-2}$, where

$m \geq 2$ is the azimuthal asymmetry index¹⁰; corresponding values for the bending-wave fit are: resonance radius $\approx 118,456$ km (see Fig. 1) and mass density $\approx 21.5 (m-1) \text{ g cm}^{-2}$. We know of no resonances with exterior satellites that can be reconciled with the two locations above. Furthermore, the two implied mass densities differ significantly from each other, despite the spatial proximity of the observed patterns and the comparable value of their background optical depth. In addition, even for $m = 2$, the mass density implied by the density wave is significantly larger than values obtained elsewhere in the rings for regions of much larger background optical depth^{11,12}. Thus, in all likelihood, the fluctuations observed are not density or bending waves.

Given the most likely location of a satellite, the complete theoretical wake profile can be computed from equations (1) and (5) for assumed values of background optical depth τ_0 and mass relative to the mass of Saturn, M_s/M_p . Several theoretical profiles are compared with the data in Fig. 4*a*; the background opacities for the models are assumed to be $\tau_0 = 0.2$ and 0.15 for the inner and outer wakes, respectively.

Comparisons similar to those in Fig. 4*a* have led to the conclusion that there is no single theoretical profile that satisfactorily fits all oscillations observed simultaneously, even when τ_0 is allowed to vary across the wake. A mass large enough to account for the amplitudes observed farthest from the gap ($M_s/M_p \approx (20-30) \times 10^{-12}$ for the inner wake and $(8-10) \times 10^{-12}$ for the outer wake) overestimates the amplitude of intermediate oscillations and implies that oscillations observed closest to the gap are shocked. On the other hand, a smaller mass sufficient to account for the amplitudes of the oscillations closest to the gap ($M_s/M_p \approx 10 \times 10^{-12}$ and $(4-6) \times 10^{-12}$ for the inner and outer wakes, respectively) leads to amplitudes that seriously underestimate those observed farthest from the gap. The limitation of non-colliding streamlines and the neglect of self-gravity of the disk appear likely reasons for this dichotomy.

In view of this model limitation, it is probably meaningful to calculate only bounds on the masses, by fitting the individual oscillations observed rather than the full wake pattern. Values

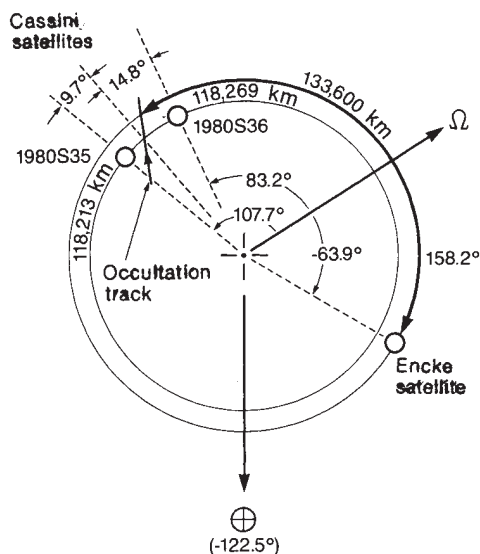


Fig. 5 Locations of the two Cassini satellites relative to the locus of intersection of the ray from Voyager to Earth with the ring plane, labelled 'Occultation track'. Longitude is measured anti-clockwise from the intersection of Earth's mean Equator (1950.0) with the ring plane at the ascending node (the radial vector labelled Ω). Also shown are the projected Earth direction ($\oplus$) and, for comparative purposes, the location of a third satellite in the Encke gap of Saturn's rings whose presence was detected by several Voyager observations, including radio occultation³⁻⁶. The epoch of all longitudes shown is $t_{\text{ref}} = 1980:318:04:54:39.39$ (UTC/ERT).

for M_s/M_p listed in Table 1 correspond to best fits to the ratio τ_{i+1}/τ_i , which eliminates dependence on the background opacity τ_0 in the model equation (5); they are also plotted in Fig. 4b against the mean radial distance, $(r_{i+1} + r_i)/2$. The general trend of smaller mass ratio for oscillations at smaller distance from the gap is evident. The approximate range for the mass ratio of the satellite responsible for the outer wake is $M_s/M_p \approx (4-8) \times 10^{-12}$ ($M_s = 2.3-4.6 \times 10^{18}$ g) and for the satellite responsible for the inner wake is $M_s/M_p \approx (1-2.5) \times 10^{-11}$ ($M_s = 5.7-14.2 \times 10^{18}$ g). Assuming a solid sphere of material density 1 g cm^{-3} , these masses correspond to spheres of radii $\sim 8-10$ km and $11-15$ km, respectively. We emphasize the preliminary nature of these mass estimates, as the linear model used does not agree with the observations regarding the amplitudes of the wake (see Fig. 4).

Discussion

Our main results are summarized in Fig. 5, which is a view in the ring plane showing the orbits of the two satellites and their longitudes relative to the radio occultation track (the locus of intersection of the Voyager-Earth line with the ring plane). The absolute longitude reference and the epoch are indicated in Fig. 5 legend. The inner and outer satellites have been given the provisional IAU designations 1980S35 and 1980S36, respectively¹³. Geometrically, they appear to shepherd a 30-km-wide ringlet within the 100-km gap in which they exist (see Figs 1, 2b). Whether they are dynamical shepherds¹⁴ remains to be seen.

For comparative purposes, Fig. 5 also shows the relative geometry of a third satellite, identified by similar techniques in the Encke gap³⁻⁶. Its mass is estimated to be $(5-10) \times 10^{-12}$ Saturn masses⁵, a value comparable to that of the inner Cassini satellite; however, the Encke gap is nearly eight times as wide (~ 322 km) as the section of the gap within which the Cassini satellite exists (~ 40 km). If these gaps are assumed to be maintained by the gravitational torques of the embedded satellites, balanced by the torque due to viscous spreading of adjacent ring material^{1,8,15,16}, then the balance would imply an unusually large dispersion velocity of 7.7 cm s^{-1} in the neighbourhood of the 40-km-wide Cassini gap, compared with only 0.2 cm s^{-1} for the Encke gap, assuming a typical mass ratio of $\sim 10^{-11}$ in both cases. However, as discussed more fully in ref. 4, a simple torque balance cannot account for observed sharp gap-edges; furthermore, it is also inconsistent with the apparent piling-up of ring material near the two edges of the Cassini gap at 118,245 km (see Fig. 1). Thus, the implications for viscosity above must be regarded with due care. Even more unusual in this regard is that the more massive outer Cassini satellite is accommodated in a narrower gap, of width comparable to the satellite's own physical size (~ 30 km).

Why were such massive satellites not detected in Voyager 2 images of the Cassini division? A careful search for embedded satellites over nearly 340° coverage of images, processed to enhance an initial resolution of ~ 80 km per line pair, proved negative². However, the search was limited to the division's two widest gaps and rules out satellites of radii ≥ 5 km if their albedo exceeds 0.1. In the relatively 'bright' environment of Saturn's rings a smaller satellite albedo may be difficult to justify, but is nevertheless a possibility. It is more likely, however, that no single physical mechanism is responsible for the existence of all gaps. Unfortunately, a similar search for satellites embedded within the other narrower gaps is hampered by resolution limitations (J. N. Cuzzi and J. J. Lissauer, personal communication).

Are the two Cassini satellites identified in the radio data the only relatively massive bodies within that division? This seems unlikely to us, given the low probability of the unusual geometry of the satellites relative to the occultation track (see Fig. 5), and given also the spatial proximity of two other gaps of width nearly equal to that of the gap sections accommodating the identified satellites (see Fig. 1). Why have no other wakes been detected, therefore, considering that the Encke satellite wake was observed $\sim 158^\circ$ away from the satellite (see Fig. 5)? The physical nature of the background media supporting the Cassini and Encke wakes is substantially different in the two cases, being optically thin and nearly devoid of centimetre-sized particles in the Cassini case, and relatively more opaque with the abundant presence of small particles in the Encke case^{3,17}. The impact of this difference on the evolution of wake profiles as a function of angular separation from the satellite, in the presence of collisions and self-gravity, and also on the width of the gap within which the satellite exists, is not yet understood.

We thank M. R. Showalter and J. N. Cuzzi for helpful discussions, P. Goldreich and J. J. Lissauer for comments on the original manuscript, R. A. Simpson and P. A. Rosen for help with the initial data reduction, and V. R. Eshleman for encouragement. This work was supported by NASA and the Voyager project.

Received 14 April; accepted 1 July 1986.

1. Lissauer, J. J., Shu, F. H. & Cuzzi, J. N. *Nature* **292**, 707-711 (1981).
2. Smith, B. A. *et al. Science* **215**, 504-537 (1982).
3. Marouf, E. A., Tyler, G. L. & Rosen, P. A. *Icarus* (in the press).
4. Cuzzi, J. N. & Scargle, J. D. *Astrophys. J.* **292**, 276-290 (1985).
5. Showalter, M. R., Cuzzi, J. N., Marouf, E. A. & Esposito, L. W. *Icarus* **66**, 297-323 (1986).
6. Marouf, E. A. & Tyler, G. L. *Bull. Am. astr. Soc.* **17**, 717 (1985).
7. Borderies, N., Goldreich, P. & Tremaine, S. *Icarus* **63**, 406-420 (1985).
8. Goldreich, P. & Tremaine, S. *A. Rev. Astr. Astrophys.* **20**, 249-283 (1982).
9. Lissauer, J. J. & Peale, S. J. *Icarus* (in the press).

10. Shu, F. H. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 513-561 (University of Arizona Press, Tucson, 1984).
11. Cuzzi, J. N. *et al. in Planetary Rings* (eds Greenberg, R. & Brahic, A.) 73-199 (University of Arizona Press, Tucson, 1984).
12. Esposito, L. W. *et al. in Saturn* (eds Gehrels, T. & Matthews, M. S.) 463-545 (University of Arizona Press, Tucson, 1984).
13. *IAU Circ. No. 4178* (1986).
14. Goldreich, P. & Tremaine, S. *Nature* **277**, 97-99 (1979).
15. Hénon, M. *Nature* **293**, 33-35 (1981).
16. Hénon, M. in *Planetary Rings* (ed. Brahic, A.) 363-384 (Centre National d'Etudes Spatiales, Toulouse, France).
17. Tyler, G. L. *et al. Icarus* **54**, 160-187 (1983).

Polymer inaccessible volume changes during opening and closing of a voltage-dependent ionic channel

Joshua Zimmerberg & V. Adrian Parsegian

Division of Computer Research and Technology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

Osmotic stress can be used to estimate the internal volume change during the opening and closing of a voltage gated ionic channel. Mitochondrial voltage-dependent anion channels, from rat liver and from Neurospora, reconstituted into planar lipid bilayers show a change of 2 to $4 \times 10^4 \text{ \AA}^3$ in internal volume, a large change inconsistent with a blocking or local gating model but supporting models with major closure of the channel space.

THE growing precision with which one can describe the electrical properties of single ionic channels and the increasing definition of membrane structure by a dozen or more probes have reinforced the common assumption that electrically excited ionic channels can be envisaged as a chemical species whose functional variations are correlated with structural modifications. The current status of this correlation is expressed in a very long list of geometric models proposed on the basis of electrical behaviour, kinetics, and incomplete structural information. To our eyes, these models divide roughly either into 'blocking' pictures, where something—such as a cork, a flap, or a local constriction—stops the channel, or into 'rearrangement' pictures, where large parts move to close down a significant fraction of the volume through which the ions move. Because these two classes of models usually predict clearly different changes in internal channel volume with channel opening and closing, we believe one might distinguish between models or even between blocking and rearrangement classes of models by measuring the changes in the internal volume inaccessible to pore impermeant species. As a first example, we have chosen to measure volume changes in the voltage-dependent anion channel (VDAC)¹ or mitochondrial porin^{2,3} from the outer membrane of mitochondria.

Measured volume changes are large. They correspond to a loss of most of the open channel volume and suggest, to us, rearrangements of the type proposed by Unwin and Ennis⁴ for gap junctions or by Doring and Colombini⁵ for VDAC rather than a gating or blocking picture.

We use a variant of the osmotic stress technique originally developed to measure forces directly between bilayer mem-

branes⁶⁻⁹ and between molecules¹⁰. Impermeant polymer of equal concentrations is introduced on both sides of a membrane that contains an excitable channel. The act of channel opening, entailing an increase in volume v ,

$$\Delta v = v_{\text{open}} - v_{\text{closed}} \quad (1)$$

incurs an extra work, $\Pi_{\text{osm}} \Delta v$, where Π_{osm} is the difference in osmotic pressure between the bathing solution and that of the pore interior, in this case exerted by the polymer in the solution (Fig. 1).

This extra work can be incorporated into the Boltzmann statistical formalism of Ehrenstein *et al.*¹¹ as adapted by Schein *et al.*¹

$$\frac{N_0}{N_c} = \exp \left[-nq(\psi - \psi_0)/kT - \Pi_{\text{osm}} \Delta v/kT \right] \quad (2)$$

Here N_0 and N_c are the times spent in the open and closed states, n is the number of imagined 'gating' elementary ionic charges q traversing the entire transmembrane voltage ψ during channel opening; ψ_0 is that applied voltage at which channels are expected to be open half the time under no added osmotic stress; and kT retains its usual meaning. Application of osmotic stress can be expected to shift the apparent ψ_0 by $\Pi_{\text{osm}} \Delta v/qn$. By measuring Π_{osm} , n and apparent ψ_0 , we infer

$$\Delta v = (qn/\Pi_{\text{osm}}) \times \text{shift in apparent } \psi_0 \quad (3)$$

Our first observations were on membranes containing many channels to determine relative membrane conductance versus applied voltage. Because of the indeterminate number of chan-

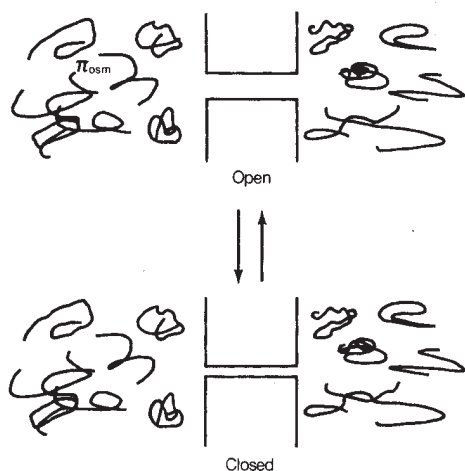
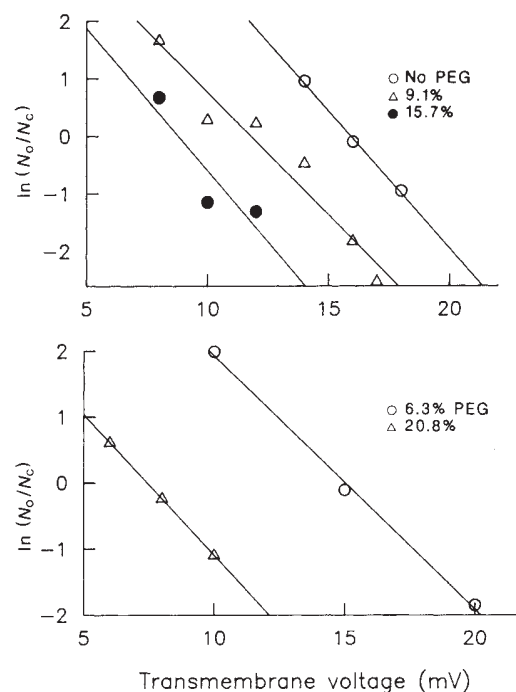


Fig. 1 Scheme of volume change measurement. Impermeant polymer on either side of membrane competes for channel water with an osmotic pressure Π_{osm} . The energy of channel opening $\Delta E(\psi, \Pi_{\text{osm}})$ determines the ratio N_0/N_c of open to closed times by $(N_0/N_c) = \exp[-\Delta E(\psi, \Pi_{\text{osm}})/kT]$. In practice we expand $\Delta E(\psi, \Pi_{\text{osm}}) = \Delta E(\psi, 0) + \Pi_{\text{osm}} \Delta v$ where $E(\psi, 0) = nq(\psi - \psi_0)$ (for $\psi > 0$) and $\Delta v \equiv \partial \pi E / \partial \Pi_{\text{osm}}|_{\Pi_{\text{osm}}=0}$. This expansion is used here assuming that the internal polymer inaccessible volume of the open and closed states varies negligibly with applied voltage and osmotic stress.

Fig. 2 a, PEG reversibly closes VDAC. A planar phospholipid bilayer was formed from asolectin monolayers on symmetric 1M KCl, 5 mM CaCl₂. Initial conductance was 6×10^{-12} S. Seven channels were observed to insert into the bilayer after addition of 0.2 μ l of Triton-solubilized outer mitochondrial membranes (rat liver) to the ground chamber. The membrane was subjected to the indicated voltages and the open/closed ratio of the channels was determined ($\circ$) by observing discrete jumps in current. After addition, with stirring, of PEG M_r 20,000 in 1M KCl, 5 mM CaCl₂ to both chambers (final concentration 15.7%) another set of data were obtained ($\bullet$). Finally, fresh 1M KCl, 5 mM CaCl₂ was perfused into both chambers (final PEG 9.1%) (Δ). **b**, Continuous current experiment. A planar phospholipid bilayer was prepared containing a large number of channels which fluctuated throughout the experiment. PEG was added at an initial level of 6.3% ($\circ$) and then 20.8% (Δ). At each polymer concentration the indicated voltage was applied across the bilayer and the steady state conductance (g) measured after a long pulse (10–30 min). After estimation of g , the transmembrane voltage was stepped to zero to turn all channels on and to correct for a small drift in the junction potentials of the calomel electrodes (<1 mV). Then the maximal conductance ($g_{\max}$) was determined repeatedly by measuring the instantaneous current response to a change in transmembrane voltage. The voltage-independent conductance ($g_{\min}$) was determined by applying a large voltage to the membrane which closed all channels. **Methods.** Crude outer mitochondrial membranes were obtained by differential centrifugation of rat liver according to the procedure of Parsons *et al.*¹². Outer membranes were sometimes purified further by NaI extraction¹³. The first outer mitochondrial membrane pellet was resuspended in 50 ml of 1 mM KCl 1 mM Hepes (N-2-hydroxyethyl piperazine-N-2 ethansulphonic acid), pH 7.5 and pelleted for 20 min at 35,000g in a SS-34 rotor in a Sorval RC-2B centrifuge. This suspension was either sonicated with an excess of asolectin, lyophilized, and suspended in *n*-pentane (1% w/v) or extracted with 2% Triton X-100 (less than 6 mg ml⁻¹ protein)¹⁴ for 30 minutes in 1 mM KCl, 1 mM, then quick-frozen and stored in liquid nitrogen after addition of 20% (v/v) dimethylsulphoxide¹⁵. We reconstituted VDAC into planar phospholipid bilayers by the method of Colombini¹⁴. We used the Montal-Mueller procedure to make planar phospholipid bilayers in lucite chambers separated by a thin partition of Teflon containing a 150 μ m hole¹⁸. We then inserted channels into the stable membrane by adding 0.1 to 0.5 μ l aliquots of the triton solubilized outer membrane solutions (to which 0.1% sonicated asolectin had been freshly added) to the membrane bath¹⁴. All aqueous phases contained 1M KCl and 5 mM CaCl₂. Current and voltage were, respectively, monitored and enforced with a single pair of calomel electrodes via saturated KCl junctions connected to an operational amplifier (42L, Analog Devices). Data were either recorded on a strip chart recorder or a computer. Asolectin bilayers typically showed high resistance before VDAC addition, approximately $10^8 \Omega \text{ cm}^2$. Thus the conductance of the unmodified bilayer did not contribute to the large, voltage-dependent conductances induced by VDAC.



nels being observed and of the finite conductance of channels in the 'closed' stage, it was necessary to look at the full range of membrane conductances from $g_{\min}$ where all channels are expected to be closed to $g_{\max}$ where all are open and to derive the open to closed ratio from¹

$$\frac{N_0}{N_c} = \frac{g_{\min} - g}{g - g_{\max}} \quad (4)$$

In several experiments the number of channels in the membrane was sufficiently small that low noise levels allowed for a single channel opening or closing transition to be detected. In these cases the open to closed ratio was directly estimated by summing the time spent in each state at a given voltage.

We apply a given transmembrane voltage, find it necessary to wait approximately 20 minutes for equilibration, then record the conductance $g(\psi)$. (The higher the potential, the sooner equilibrium is reached.) After a number of such measurements at different voltages, this process is repeated with increasing concentrations of polymer to ascertain that there is in fact a shift in $g(\psi)$, indicating the action of osmotic stress. Three different polymers were used in different experiments to ensure that their action was indeed a function only of osmotic pressure and not due to specific binding to the channel, viscosity, presence of a contaminant, or some other property of the test solution. On three occasions polymer was washed out to verify recovery of ψ_0 and n . Polymer concentration in the aqueous phase was measured directly from the viscosity of small withdrawn samples. Because n and ψ_0 vary between preparations¹, we use differences in apparent ψ_0 only within the same planar bilayer. We used VDAC from rat liver and from *Neurospora crassa* mitochondria.

Polymer closes VDAC

Adding polymer to both sides of a membrane containing VDAC led to a reversible downward shift in the conductance-voltage curve. We first measured the membrane conductance at different voltages in the absence of polymer. In all experiments channels were always open at sufficiently low voltage, but closed at higher voltages as expected from equation (2). After adding polymer we found significant channel closure at lower transmembrane potentials, including those at which the channels had previously always been open.

The downward shift is clearly seen in the experiments of Fig. 2. In Fig. 2a a concentration of 15.7% polyethyleneglycol (PEG- $2.84 \times 10^6 \text{ dyn cm}^{-2}$) causes a bias towards the closed state at any potential. After removal of some of the polymer, to 9.1% PEG ($8.37 \times 10^5 \text{ dyn cm}^{-2}$), this bias is decreased.

In other experiments, virtually complete removal of polymer resulted in recovery of the original conductance-voltage relationship. The slope of the voltage dependence (usually thought to be related to n , the number of gating particles) did not change consistently as a function of polymer concentration. The effect of added polymer was symmetric with respect to the sign of the applied voltage.

Qualitatively similar results were observed in membranes containing a single conducting unit (Fig. 3). Because of the slow kinetics, the number of transitions is necessarily too small to provide reliable statistics. The most important lesson from these single channel measurements is that polymer alters neither the predominant open and closed state conductances nor the general appearance of channel opening and closing. Data from an oligochannel experiment (Fig. 2a) were analysed to determine the effect of polymer on single-channel conductances. A

Table 1 Determination of volume changes on planar membranes with few (D = discrete) or very many (C = continuous) VDAC channels

Polymer	Type	$\Delta\Pi_{\text{osm}}$ (dyn cm ⁻²)	(Π_1 to Π_2)	$\Delta\psi_0$ (mV)	n	Δv (Å ³)
PEG	D	2.9×10^6	(0, 2.9)	7.1	12, 13	4.8×10^4
PEG	D, W	2.1×10^6	(2.9, 0.8)	3.0	13, 12	2.8×10^4
PEG	C	5.3×10^6	(0.4, 5.7)	7.4	10, 11	2.3×10^4
PEG	C, W	2.4×10^6	(3.0, 0.6)	2.8	11, 15	2.4×10^4
PEG	C	3.5×10^6	(1, 4.5)	5.2	10, 11	2.6×10^4
PEG	C, N	1.1×10^6	(0.3, 1.4)	1.8	13, 11	2.2×10^4

Channel types: D, discrete; C, continuous; W, wash out; N, *Neurospora*. After determination of n and ψ_0 , we stirred polymer into the solutions bathing both sides of the planar membrane while being careful to preserve electrolyte activities. The apparent n and ψ_0 were determined repeatedly as a function of increasing polymer concentration. The shift in ψ_0 was determined for each sequential step in osmotic stress from Π_1 to Π_2 . Rat liver VDAC was used unless *Neurospora* (N) is specified. Partially purified *Neurospora* VDAC was a generous gift of C. Mannello^{16,17}. The stressing polymers were polyethyleneglycol (PEG) of M_r 20,000 from Fluka (Hauppauge), polyvinylpyrrolidone (PVP) M_r 40,000 from Aldrich, and Dextran 1500 (M_r 500,000) from Pharmacia Fine Chemicals. The dextran was dialysed against distilled water extensively before use. Polymer osmotic pressures were measured by N. Fuller (dextran and PEG) and D. Rau (PEG and PVP). Asolectin was obtained after ether/acetone extraction of neutral lipids (Avanti).

difference in single channel conductance of 2.54 ± 0.29 nS between open and closed states is found without PEG; 2.4 ± 0.3 nS with 15.7% PEG.

Data from a set of measurements such as those in Fig. 2 permit us to make several independent calculations of the change in the volume of the polymer-inaccessible space during opening and closing. The shifts in apparent ψ_0 with changes in applied pressure are summarized in Table 1. A volume change of approximately 3×10^4 Å³ could account for these shifts.

Additional experiments, in which more than 10 transitions from open to closed were observed per voltage point and in which the open to closed ratio was determined at two osmotic stresses with the same voltage, are summarized in Table 2. In this case the electrical energy does not change and

$$\Delta v = kT(\Delta \ln N_o/N_c)/\Delta\Pi_{\text{osm}} \quad (5)$$

The point of this table is to indicate that we find no systematic variation of Δv with voltage, consistent with our original assumption. This is equivalent to seeing no change in n .

To ensure that the polymer action was indeed a function only of osmotic pressure and not due to specific binding to the channel, viscosity, presence of a contaminant, or some other property of the test solution, we have used three different polymers in different experiments with similar results (Tables 1, 2). For the same polymer, the results were largely independent of concentration, a result incompatible with polymer binding to the channel or some 'pharmacological effect'. The small variation seen in volume estimates may reflect channel elasticity under stress. We see no obvious dependence of estimated volume change on the molecular weight of the polymer.

To control for variations in water activity and hydrogen bonding to the channel itself, we used three saccharides at comparatively high concentrations to look for a shift in ψ_0 . Glucose, sucrose, and stachyose were added to the buffer bathing bilayers containing VDAC to increment osmotic pressure by 1.5×10^7 , 1.3×10^7 , and 6.7×10^6 dyn cm⁻², respectively. No change in channel open/closed statistics were seen. Thus permeant solutes of the same chemical species as the larger polymer dextran could not bias channel rearrangement.

Many times the polymer seemed to induce a conductance state not seen in its absence. In one case we calculated a volume change of the substate by $\Delta(N_o/N_c)/\Delta\Pi_{\text{osm}}$ of $7,900$ Å³. A small

change in the open conductance, seen in the Fig. 3 histograms (arrow), was seen with polymer and again appeared to be a conductance state that is electrically inaccessible in the absence of polymer.

Discussion

The 2.2 to 4.8×10^4 Å³ volume change which we measure compares reasonably with estimates of open channel volume from two methods. First, the radius of the VDAC pore was determined to be 20 Å by permeability studies¹⁹, but only some 9 Å from single channel conductance²⁰. Electron microscopy and reconstruction of fungal VDAC arrays gives radii of 10 to 15 Å (ref. 21). Similar channel arrays have been seen on negatively stained rat liver outer mitochondrial membranes. In side view these channels have the same inner diameter as fungal VDAC and are at least 60 Å long. This length and the 10 to 20 Å range of radii suggest a channel volume range of 2 to 8×10^4 Å³. The volume change we measure is a large fraction of the total volume of VDAC even though 'closure' is not complete. This suggests that channel closure incurs a relatively large loss of channel volume rather than the local blockage or rotation of a single group frequently associated with the gating process. Of the currently available models of channel opening and closing, those of Doring and Colombini⁵ and Unwin and Ennis⁴ seem to us to come closest to the kind of picture required, at least by VDAC.

The large volume change observed and the method used to detect it suggest that the water inside a channel is probably subject to the same kind of perturbation that causes the hydration forces seen between membranes^{6,8} and macromolecules¹⁰. Channel opening and closing could involve hydration/dehydration energetics whose magnitude depends on the species of ion in the channel much as intermolecular hydration forces between macromolecules change with the identity of associated counterions¹⁰.

We should think of the channel protein as a chemical species whose various states are functions of a potentially large family of system variables. Osmotic stress and associated water are but one pair of conjugate variables. Others include hydrostatic pressure and specific molecular volume (density); the volume occupied in the membrane (outer diameter) and membrane rigidity. We note that these pairs of variables are not *a priori* coupled to $\Pi, \Delta v$. In particular, channel gating changes in response to hydrostatic pressure²² are not necessarily relevant to pore volume changes. The conformational changes that we see in membrane proteins under osmotic stress may also be expected in other macromolecules whose different states require different states of hydration. For example, enzyme activities and

Table 2 Δv measured at constant voltage

Polymer	Type	$\Delta\Pi_{\text{osm}}$ (dyn cm ⁻²)	(Π_1, Π_2)	Δv (Å ³)
Dextran	C, F	9.8×10^5	$(0, 9.8) \times 10^5$	3.8×10^4
PVP	D	5.2×10^5	$(12.7, 17.9) \times 10^5$	2.6×10^4
PVP	D	4.8×10^5	$(2.6, 7.4) \times 10^5$	3.9×10^4
PEG	C, N	1.0×10^6	$(20, 30) \times 10^5$	4.1×10^4
PEG	C, N	7.9×10^5	$(6.3, 14.2) \times 10^5$	2.9×10^4
PEG*	S	1.1×10^7	$(0, 110) \times 10^5$	3.82×10^4

C, continuous; D, discrete; S, single channel; N, *Neurospora*; F, folded membrane reconstitution. The folded membrane reconstitution was a modification of the technique of Schein *et al.*¹. The pentane suspension was sonicated under argon in a bath sonicator (Cole-parmer) for 5 minutes. We layered $5 \mu\text{l}$ of the mixture into aqueous solutions to yield lipid-protein monolayers which were then used to form bilayers containing VDAC as previously described¹. Methods and abbreviations as in Table 1 except for the folded membrane reconstitution. The volume change was calculated from the shift in N_o/N_c at a single voltage, thus setting $(\psi - \psi_0)$ the same with and without polymer.

* Assuming $v_0 = 15$, $n = 11$ (mean for data set of rat liver mitochondria).

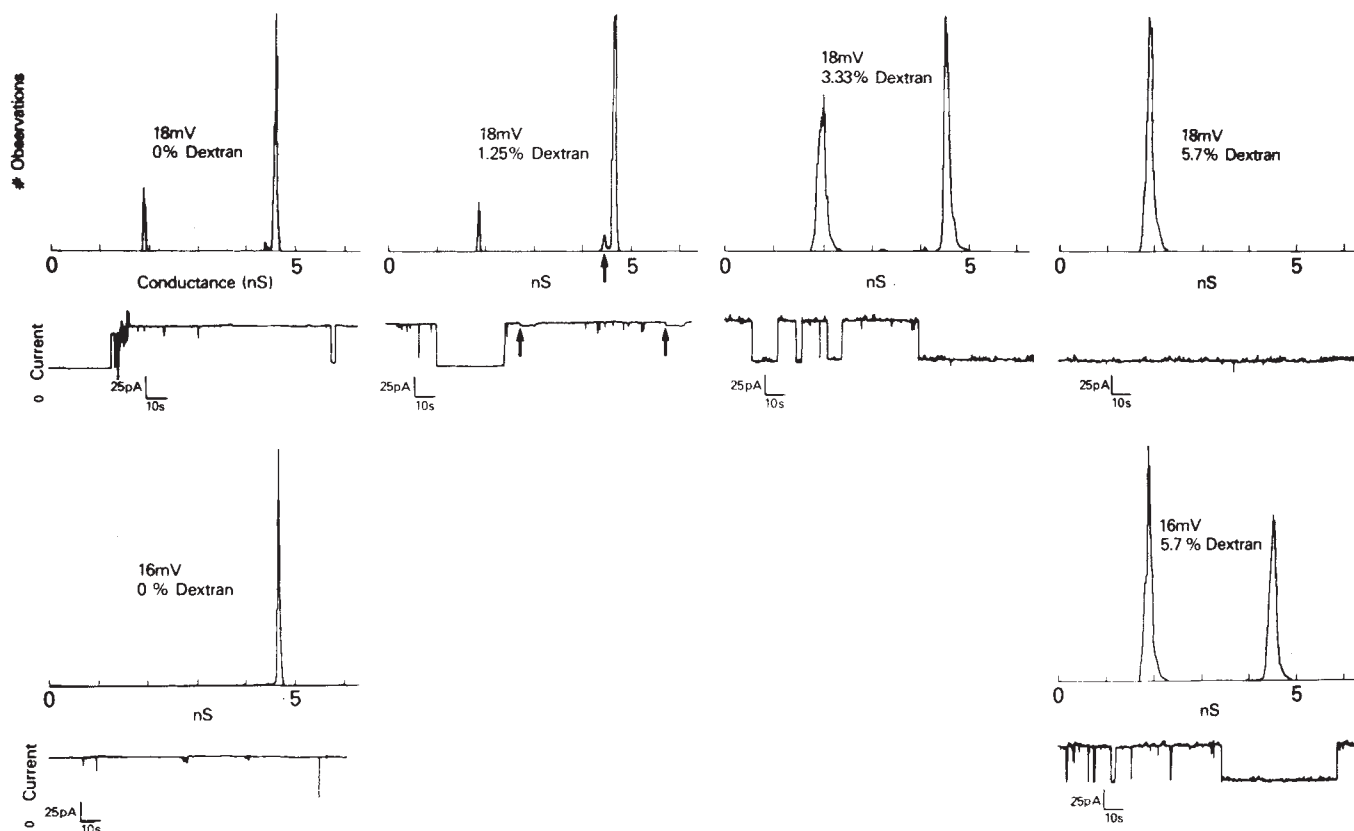


Fig. 3 Dextran decreases single channel open time. A membrane was subjected to 16 and 18 mV transmembrane potential at various dextran concentrations. Histograms of conductance levels were constructed as a function of applied voltage in order to determine the fraction of time channels were in 'open' and 'closed' states. Initially the channel was completely open at 16 mV and mostly open at 18 mV (left side of figure). Subsequent addition of dextran led to predominant closure at 16 mV and complete closure at 18 mV. Representative traces show no change in the (large step) single channel conductance. In this run, for an asolectin planar phospholipid bilayer in 1M KCl, one channel entered the bilayer five minutes after addition of 0.1 μ l solubilized VDAC. Other channels inserted after collection of the data shown. The dextran suppression of opening was reversed by dextran washout at this later stage. An upward transition in current denotes channel opening. A possible open conductance substate was seen with polymer addition (arrow). A portion of the current trace on the chart recorder is shown for illustration below each histogram.

allosteric mechanisms which alter solute inaccessible water spaces will be biased by osmotic pressure. For the opening and closing of very large channels such as VDAC, the prevailing intracellular osmotic pressure can act in tandem with transmembrane potential to control channel conductance, suggesting a role for VDAC in sensing or controlling intracellular colloidal pressure. Further, the study *in vitro* of any system in the absence of normal cellular osmotic stress might systematically bias our view of that system.

We are now applying osmotic stress to measure volume

changes in several other channel systems²³. For smaller pores one must apply higher pressures (and can use smaller solutes) to reach conditions where $\Pi_{\text{osm}}\Delta v$ is the order of kT where conductance effects are expected.

We thank Gerald Ehrenstein for the use of laboratory facilities; Kuni Iwasa for advice on channel analysis software; Donald Brooks for suggesting that we dialyse our polymer solutions; Marco Colombini for occasional use of his laboratory plus practical advice; and Carmen Mannella for a gift of *Neurospora* VDAC.

Received 3 February; accepted 3 July 1986.

- Schein, S. J., Colombini, M. & Finkelstein, A. *J. Membrane Biol.* **30**, 99-120 (1976).
- Roos, N., Benz, R. & Brdiczka, D. *Biophys. biochim. Acta* **686**, 204-214 (1982).
- Benz, R. *CRC crit. Rev. Biochem.* **19**, 145-190 (1985).
- Unwin, P. N. T. & Ennis, P. D. *Nature* **307**, 609-613 (1984).
- Döring, C. & Colombini, M. *J. Membrane Biol.* **83**, 87-94 (1985).
- LeNeveu, D. M., Rand, R. P. & Parsegian, V. A. *Nature* **259**, 601-603 (1976).
- Rand, R. P. *A. Rev. Biophys. Bioeng.* **10**, 277-314 (1981).
- Lis, L. J., McAlister, M., Fuller, N., Rand, R. P. & Parsegian, V. A. *Biophys. J.* **37**, 657-666 (1982).
- Parsegian, V. A., Rand, R. P. & Rau, D. C. *Meth. Enzym.* **127**, 400-416 (1986).
- Rau, D. C., Lee, B. & Parsegian, V. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2621-2625 (1984).
- Ehrenstein, G., Lecar, H. & Nossal, R. *J. gen. Physiol.* **55**, 119-133 (1970).

- Parsons, D. F., Williams, G. R. & Chance, B. *Ann. N.Y. Acad. Sci.* **137**, 643-666 (1966).
- Cohen, F. S., Zimmerberg, J. & Finkelstein, A. *J. gen. Physiol.* **75**, 251-270 (1980).
- Colombini, M. *Nature* **279**, 643-645 (1979).
- Colombini, M. *J. Membrane Biol.* **74**, 115-121 (1983).
- Benz, R., Ludwig, O., de Pinto, V. & Palmieri, F. in *Achievements and Perspectives of Mitochondrial Research Vol. 1* (eds Quagliariello, E., Slater, E. C., Saccone, C., Palmieri, F. & Kroon, A. M.) 317-327 (Elsevier, Amsterdam, 1985).
- Mannella, C. A. *J. Cell Biol.* **94**, 680-687 (1982).
- Mannella, C. A. *Science* **224**, 165-166 (1984).
- Montal, M. & Mueller, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3561-3566 (1984).
- Colombini, M. *Ann. N.Y. Acad. Sci.* **341**, 551-563 (1980).
- Benz, R., Ludwig, O., de Pinto, V. & Palmieri, F. in *Achievements and Perspectives of Mitochondrial Research Vol. 1* (eds Quagliariello, E., Slater, E. C., Saccone, C., Palmieri, F. & Kroon, A. M.) 317-327 (Elsevier, Amsterdam, 1985).
- Mannella, C. A., Radermacher, M. & Frank, J. *Proc. 42nd Annual Meeting of the Electron Microscopy Society of America*, 644 (1984).
- Conti, F., Fioravanti, R., Segal, J. R. & Stühmer, W. *J. Membrane Biol.* **69**, 35-40 (1982).
- Benzanilla, F., Parsegian, V. A. & Zimmerberg, J. *Biophys. J.* **49**, 161a (1986).

Cosmic strings, superstrings and the evolution of the Universe

L. J. Tassie

Department of Theoretical Physics, Research School of Physical Sciences, The Australian National University, Canberra, Australia

For a wide variety of astronomical objects, the angular momentum is proportional to the square of the mass, and the constant of proportionality is comparable with that of string theories of particle physics, suggesting that the Universe has evolved through a hierarchical breaking of rotating pieces of string. I show here that if such a picture is correct, the more distant objects in the Universe should appear more string-like.

For astronomical objects of different classes, ranging from planets to superclusters, Brosche^{1,2} and Wesson^{3,4} have claimed that the variation of angular momentum J with mass M between the various classes is given by

$$J = \kappa M^2 \quad (1)$$

where κ is a universal constant, although within each class there is a considerable scatter of the data in the J - M plane and the best fit to the J - M relationship may be $J = kM^\beta$, with $\beta \neq 2$, where k depends on the class of object. An indication of the evidence for equation (1) is given in Table 1. As discussed in ref. 4, this evidence is most readily seen in the compilation of the data given by Carrasco *et al.*⁵, who find $\frac{5}{3} \leq \beta \leq \frac{7}{4}$ within each class of object, but for all classes of objects find $\beta = 1.94 \pm 0.09$.

For all classes of object, equation (1) gives a good description of the general trend of the data over 24 powers of ten in the mass, and so allows a determination of κ . Wesson's analysis³ gave the value $\kappa = 8 \times 10^{-16} \text{ g}^{-1} \text{ cm}^2 \text{ s}^{-1}$, although there is considerable uncertainty in the value of κ , as Wesson obtained $\log_{10} \kappa = -15.1 \pm 0.9$. Wesson's value for κ can be written as $\kappa = 4 \times 10^2 G/c$ where G is the universal gravitational constant and c is the speed of light.

Wesson^{3,6} has given a dimensional argument for equation (1), but did not provide any physical model of how this occurs. Wesson has also pointed out that the value obtained for κ is consistent with $\kappa = G/\alpha$ where α is the dimensionless electromagnetic fine-structure constant $\alpha = e^2/4\pi\hbar c \approx \frac{1}{137}$. As remarked by Brosche², a similar relation occurs for the Regge trajectories of hadrons, for which⁷

$$J(M) = \kappa_h M^2 + J(0) \quad (2)$$

where $\kappa_h = \hbar/(\text{GeV}/c)^2 = 1.5 \times 10^{38} G/c$.

As κ_h differs so greatly from κ for astronomical objects, the spectrum of astronomical objects cannot be directly related to the spectrum of hadrons. However, there is a simple physical model of the hadrons in which the spectrum (2) is the spectrum of the rotational states of a string^{8,9}. The string parameter is μ , the mass per unit length, also called the string tension. For a rigidly rotating, straight, open string,

$$\kappa = c(2\pi\mu)^{-1} \quad (3)$$

For rigidly rotating closed strings, the right-hand side of equation (3) contains an additional factor ranging from $\frac{1}{2}$ to $\frac{1}{6}$, according to the particular class of solutions¹⁰, and such a factor can be neglected in view of the uncertainty in κ for astronomical objects.

Assuming that equation (1) for astronomical objects also represents the spectrum of a rotating string, the astronomical string has $\mu_a = 6 \times 10^{24} \text{ g cm}^{-1} = 4 \times 10^{-4} c^2/G$, whereas the hadronic string has $\mu_h = 10^{-11} \text{ g cm}^{-1} = 10^{-39} c^2/G$.

Strings can occur as vortices in gauge theories^{11,12}, in an analogous manner to the formation of Abrikosov magnetic filaments in superconductors. Kibble¹³ has suggested that such

Table 1 The range of J/M^2 for the planets³⁴, excluding Mercury, Venus and Pluto, and for the 23 flat galaxies of the compilation of Miyamoto *et al.*³⁵

	M	J/M^2 ($10^{-15} \text{ g}^{-1} \text{ cm}^2 \text{ s}^{-1}$)
Planet-moon(s) system		
Earth	$6.0 \times 10^{27} \text{ g}$	9.4
Jupiter	$1.9 \times 10^{30} \text{ g}$	1.2
Galaxy		
NGC3109	$3.6 \times 10^{43} \text{ g}$	4.7
NGC4826	$2.8 \times 10^{43} \text{ g}$	1.1

Data are from an unpublished report by T. J. Allen.

vortex strings could be formed in a phase transition in the very early stages of the Universe¹⁴. For vortex strings in a spontaneously broken gauge theory with a Higgs field^{12,15} (using units with $\hbar = c = 1$):

$$\mu = \frac{1}{4} \frac{M_x^2}{\alpha_G} f_G = \pi \eta^2 f_G \quad (4)$$

where η is the vacuum expectation value of the Higgs field, M_x is the mass of the associated vector boson, $\alpha_G = g^2/4\pi$ is the gauge field coupling constant and f_G is a geometrical factor depending on the details of the breaking of the gauge symmetry group. For the abelian Higgs model, if $M_H = M_x$, where M_H is the mass of the scalar Higgs particle, then¹⁵ $f_G = 1$.

For electro-weak strings with $\alpha = 10^{-2}$ and $M_x \approx 100 \text{ GeV}$, $\mu_{ew} \approx 2 \times 10^{-5} \text{ g cm}^{-1} \approx 2 \times 10^{-33} c^2/G$. For grand unification strings, Vilenkin¹⁶ takes $\alpha_G \approx 10^{-2}$ and $M_x \approx 10^{15} \text{ GeV}$, yielding $\mu_{GUTs} \approx 2 \times 10^{21} \text{ g cm}^{-1} \approx 2 \times 10^{-7} c^2/G$. However, Kibble¹³ points out that this value can vary considerably within the context of grand unified theories (GUTs). The value of μ for the astronomical string is much closer to that of the GUTs string than to that of either the hadronic or electro-weak string, but the difference still seems significant. However, the experimental lower limit on baryon decay¹⁷ indicates¹⁸ that $M_x \geq 10^{15} \text{ GeV}$. Nanopoulos¹⁷ concludes that a value of $M_x \approx 10^{16} \text{ GeV}$ obtained from supersymmetric (SUSY) GUTs is consistent with the experimental lower limit on baryon decay. With this latter value of M_x , $\mu_{\text{SUSY GUTs}} \approx 10^{24} \text{ g cm}^{-1} \approx 10^4 c^2/G$, which is in as close agreement with the value for the astronomical string as can reasonably be expected without including the details of the gauge group and the details of its symmetry breaking, and considering the uncertainty in the determination of the parameters.

Kibble *et al.*¹⁹ have pointed out that strings do not occur in the simplest SU(5) grand unified theory, and have shown that strings appear in the breaking of SO(10) symmetry to SU(5) $\times$ Z_2 with $M_x \approx 10^{16}$ - 10^{17} GeV . Further symmetry breaking, which occurs in this model with a boson mass of $\sim 10^{14} \text{ GeV}$, does not affect the strings. For this model $\mu_{\text{SO}(10)}$ is $\sim 10^{-3}$ - $10^{-5} c^2/G$, which is consistent with the astronomical value of μ .

Scherk and Schwarz²⁰ described non-hadrons by a string theory in which the string parameter is related to the universal gravitational constant by interpreting the massless spin-2 state of the string as the graviton. In this model, the slope of the boson trajectories is given by

$$\kappa_{\text{nonh}} = \frac{16\pi^3 G}{g_{ew}^4} = \frac{\pi G}{\alpha^2} \sin^4 \theta_W \quad (5)$$

where g_{ew} is the electro-weak charge and θ_W is the Weinberg angle. Equation (5) yields $\kappa_{\text{nonh}} = 3 \times 10^3 G/c$, which is consistent with Wesson's determination of κ for astronomical objects.

The string of Scherk and Schwarz²⁰, combining the electro-weak interaction and the gravitational interaction, is a forerunner of the superstring²¹, which is claimed to be a theory of all

interactions, including the strong interactions of hadrons. For type I string theories^{22,23}, which describe the dynamics of open strings,

$$\mu \propto \frac{G^{1/2}}{g^2} \quad (6)$$

in 10-dimensional space-time. Assuming that six dimensions are compacted^{22,23} by imposing periodic boundary conditions with period $2\pi R$,

$$\mu_1 \propto \frac{G^{1/2}}{g^2} R^{-3} \quad (7)$$

and comparison with the astronomical value of μ is not possible without an estimate of the compaction radius R .

For the heterotic string^{24,25}, the string tension is given by²⁵

$$\mu_{II} = \frac{g^2}{32\pi^2 G} \quad (8)$$

and this relation is not affected by the compaction from 10- to 4-dimensional space-time. This will give agreement with the astronomical value for μ if $g^2 \approx 0.1$. Until the phenomenology of low-energy approximation to the heterotic string is well established, it cannot be decided if this value of g^2 is reasonable.

If it is assumed that the compaction radius is approximately the same as the length defined by μ , as has been argued by Kaplunovsky²⁶, then equation (7) for the type I string yields

$$\mu_1 \propto \frac{g^2}{G} \quad (9)$$

which is of the same form as equation (8) for the heterotic string. Assuming that R is the Planck length ($(\hbar G/c^3)^{1/2} \approx 10^{-33}$ cm), equation (9) yields

$$\mu_1 \propto g^{-2} G^{-1} \quad (10)$$

which, unless g^2 is large, may conflict with the condition $\mu G < \frac{1}{4}$ required²⁷ for the string to exist in a large external space-time. For the heterotic string, $\mu G < \frac{1}{4}$ is less restrictive, yielding

$$g^2 < 8\pi^2 \quad (11)$$

Kibble¹³ and others^{14,16,19} have suggested that GUT strings play an important part in the evolution of the Universe and that strings provide the inhomogeneity leading to the formation of galaxies. In these theories, the strings constitute only a small fraction of the mass of the Universe, galaxies are formed by the accretion of ordinary matter about the strings, and some strings may still be present, and possibly nearby.

However, the explanation of equation (1) in terms of strings suggests a very different evolution of the Universe; namely, that at some time in the past strings constituted all or nearly all of the mass of the Universe, and that astronomical objects are originally formed from string in such a way that a large piece of string, which eventually corresponds to a supercluster of galaxies, breaks into smaller pieces, which eventually correspond to clusters of galaxies. These pieces of string, in turn, break into smaller pieces corresponding to galaxies, and so on for all astronomical objects obeying equation (1). At each stage in this hierarchy of string breaking, the new pieces of string may have some vibrational energy, but the vibrational energy would have to be large compared with the mass to cause an appreciable deviation from equation (1), and also it is expected that the transfer of such vibrational energy into kinetic energy of the neighbouring strings can occur more readily than the transfer of angular momentum, so that it is expected that each piece of string would soon be in a state with the minimum energy for a given angular momentum. A similar situation occurs in nuclear physics, where an excited nucleus arising from a heavy ion collision will soon settle into a 'yrast' state, a state of maximum spin for a given energy.

A less qualitative description of string fragmentation can be obtained by modifying the treatment by Aharonov and Casher²⁸ to include the effect of angular momentum. Neglecting rotation, they showed that fragmentation of one string into strings of finite mass requires a decrease in entropy and so will not occur. The treatment here is of an open string (instead of the closed string considered by Aharonov and Casher) with large angular momentum, breaking into N open strings. N is large enough that the effect of conservation of total angular momentum on the entropy of the system of N strings is negligible. The density of states for each of the N strings is²⁸

$$\rho(m) = km^{-a} \exp(m/T_c) \quad (12)$$

where m is the mass of the string and T_c is the Hagedorn temperature of the string. For the open bosonic string in D transverse dimensions, $T_c = (3\mu/\pi D)^{1/2}$. The density of states of the initial string with angular momentum J is given by^{29,30}

$$\bar{\rho}(M, J) \approx \bar{k} M^{-a-2} \exp[MT_c^{-1} - \pi\mu J(MT_c)^{-1}], \quad J < M^2/2\pi\mu \quad (13a)$$

$$= 0, \quad J > M^2/2\pi\mu \quad (13b)$$

The approximation $J \gg (3/2)^{1/2} \pi(M^2/2\pi\mu)^{1/2}$ has been used to simplify equation (13a). Following the procedure of Aharonov and Casher²⁸, the change in entropy for the fragmentation is

$$\Delta S \approx \pi\mu J(MT_c)^{-1} - N\{T_p T_c^{-1} - \ln(p^d v)\} \quad (14)$$

where d is the number of space dimensions, and for the resulting strings NT_p is the total kinetic energy, p is the average of the absolute value of the momentum of a string, and v is the average volume occupied by the centre of mass of a string. For small J the conclusions of Aharonov and Casher apply and the fragmentation does not occur. For large J and large T_p , the entropy increases and fragmentation occurs if $\pi\mu JM^{-1} > NT_p$. Then, in order that the system of N strings does not collapse back into a single string:

$$\pi\mu JM^{-1} \geq 2GMm/R \quad (15)$$

where R is the extension of the original string. For the rotating straight string⁸, $R = M/\pi\mu$ and the condition (15) becomes $(4\pi\mu)^{-1} \geq G$, which is satisfied by the astronomical string.

Eventually, the pieces of string transform, either by a phase transition, or by some form of rapid break-up, into ordinary matter, and become the stars and planets that we now observe. The resulting objects will have much the same relation between mass and angular momentum as the pieces of string from which they evolve, as they will not lose a large proportion of their angular momentum or mass. They may radiate a large amount of energy, but this will, in general, have a small effect on the order of magnitude of their mass. Thus, the spectrum of the string would be preserved in the currently observed astronomical objects, providing an explanation of the agreement of the astronomical value of κ with that of the strings in the theory of high-energy physics.

Deviations from equation (1) must be accounted for by later evolutionary effects, such as changes in J by tidal interactions³¹ and possibly changes of M and J by collisions. For instance, it seems that Mercury, with $J/M^2 = 0.014 \times 10^{-15} \text{ g}^{-1} \text{ cm}^2 \text{ s}^{-1}$, and Venus, with $J/M^2 = 0.008 \times 10^{-15} \text{ g}^{-1} \text{ cm}^2 \text{ s}^{-1}$, have lost angular momentum. In the theory of the formation of galaxies by gravitational instability^{31,32}, the angular momenta of the galaxies is described by the tidal torque theory developed by Peebles³¹ and others³³. If the galaxies are formed from rotating strings, only the deviations from equation (1) can be due to the tidal interactions.

A consequence of the picture of the Universe presented here is that the more distant objects in the Universe should appear

to us to trace out the configurations of the strings more closely than close objects, whereas if galaxies and so on have formed by some accumulation about strings, the more distant objects should appear less string-like.

Received 16 December 1985; accepted 9 June 1986.

1. Brosche, P. Z. *Astrophys.* **57**, 143–155 (1963).
2. Brosche, P. in *Cosmology and Gravitation (Spin, Torsion, Rotation and Supergravity)* (ed. Bergmann, P. G. & De Sabbata, V.) 375–382 (Plenum, New York, 1980).
3. Wesson, P. S. *Phys. Rev. D* **23**, 1730–1734 (1981).
4. Wesson, P. S. *Astr. Astrophys.* **119**, 313–314 (1983).
5. Carrasco, L., Roth, M. & Serrano, A. *Astr. Astrophys.* **106**, 89–93 (1982).
6. Wesson, P. S. *Space Sci. Rev.* **39**, 153–161 (1984).
7. Barger, V. & Cline, D. *Phenomenological Theories of High Energy Scattering* (Benjamin, New York, 1969).
8. Marinov, M. S. *Usp. Fiz. Nauk* **121**, 377–425 (1977); *Soviet Phys. Usp.* **20**, 179–208 (1977).
9. Olesen, P. *Proc. IX Balaton Symp. on Particle Physics*, 99–114 (Balatonfured, Budapest, 1975).
10. Burden, C. J. & Tassie, L. J. *Aust. J. Phys.* **37**, 1–7 (1984).
11. Tassie, L. J. *Phys. Lett.* **46B**, 397–398 (1973).
12. Nielsen, H. B. & Olesen, P. *Nucl. Phys.* **B61**, 45–61 (1973).
13. Kibble, T. W. B. *Nucl. Phys.* **B252**, 227–244 (1985).
14. Vilenkin, A. *Phys. Rep.* **121**, 263–315 (1985).
15. de Vega, H. J. & Schaposnik, F. A. *Phys. Rev. D* **14**, 1100–1106 (1976).
16. Vilenkin, A. *Phys. Rev. Lett.* **46**, 1169–1172, 1496(E) (1981).
17. Nanopoulos, D. V. in *Proc. 12th int. Conf. on High Energy Physics*, Leipzig, Vol. 2, 36–65 (Akademie der Wissenschaften der DDR, Zeuthen, 1984).
18. Koshiha, M. in *Proc. 12th int. Conf. on High Energy Physics*, Leipzig, Vol. 2, 67–82 (Akademie der Wissenschaften der DDR, Zeuthen, 1984).
19. Kibble, T. W. B., Lazarides, G. & Shafi, Q. *Phys. Lett.* **113B**, 237–239 (1982).
20. Scherk, J. & Schwarz, J. H. *Nucl. Phys.* **B81**, 118–144 (1974).
21. Green, M. B. *Nature* **314**, 409–414 (1985).
22. Schwarz, J. H. *Phys. Rep.* **89**, 223–322 (1982).
23. Green, M. B., Schwarz, J. H. & West, P. C. *Nucl. Phys.* **B254**, 327–348 (1985).
24. Gross, D. J., Harvey, J. A., Martinec, E. & Rohm, R. *Phys. Rev. Lett.* **54**, 502–505 (1985).
25. Gross, D. J., Harvey, J. A., Martinec, E. & Rohm, R. *Nucl. Phys.* **B267**, 365–414 (1986).
26. Kaplunovsky, V. S. *Phys. Rev. Lett.* **55**, 1036–1038 (1985).
27. Gott, J. R. *Astrophys. J.* **288**, 422–427 (1985).
28. Aharonov, Y. & Casher, A. *Phys. Lett.* **166B**, 289–291 (1986).
29. Chiu, C. B. & Heilmann, R. L. *Phys. Rev. D* **4**, 3184–3201 (1971).
30. Nahm, W. *Nucl. Phys.* **B120**, 125–146 (1977).
31. Peebles, P. J. E. *The Large-Scale Structure of the Universe* (Princeton University Press, 1980).
32. Efstathiou, G. & Jones, B. J. T. *Comments Astrophys.* **8**, 169–176 (1980).
33. Fall, S. M. *IAU Symp.* No. 100, 391–399 (1983).
34. Allen, S. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
35. Miyamoto, M., Satoh, C. & Ohashi, M. *Astrophys. Space Sci.* **67**, 147–171 (1980).

Cosmic string, hydrodynamics and microanisotropies in the cosmic background radiation

S. T. Chase

Blackett Laboratory, Imperial College, London SW7 2AZ, UK

'Cosmic string' formed at the mass scale of GUT (grand unified theory) symmetry breaking ($M_x \approx 2 \times 10^{15}$ GeV) is typified by a mass per unit length $\mu G/c^2 \approx 2 \times 10^{-6}$ in dimensionless units (where G is the universal gravitational constant and c is the speed of light), corresponding to $\mu \approx 2.6 \times 10^{21}$ kg m $^{-1} \approx 4 \times 10^7 M_\odot$ pc $^{-1}$ (where $M_\odot$ is the mass of the Sun). It has been conjectured that loops of such cosmic string, presumably with radii in the range 1–10 kpc, might provide the seed perturbations which could initiate galaxy formation. Here I demonstrate the possibility of detecting such loops by observing the distortions in the cosmic background radiation (CBR) induced by the hydrodynamic interaction of isolated loops of cosmic string with the cosmic fluid, during the epoch immediately preceding recombination. These distortions occur on angular scales of a few arc minutes, and may be as large as $\Delta T/T \approx 10^{-3}$, but the areal density of kiloparsec-sized loops is anticipated to be only ~ 10 per square degree in most cosmologies, so one would not expect limited-area searches for CBR microanisotropies^{4,5} to have found them yet.

Previous work concerning the possible generation of CBR microanisotropies by cosmic string has concentrated either on the statistical properties of the energy-density perturbations due to a network of strings before recombination⁶, or on the gravitational scattering of CBR photons by very long strings in the

post-recombination era⁷. In both cases it is concluded that root-mean-square fluctuation values are well below current observational limits on all scales. The hydrodynamic effect of small isolated loops of string before recombination has not so far been considered.

As the size of the horizon at recombination ($\lambda_h \approx 100$ –200 kpc) is much greater than the size of the loops considered here ($R \approx 0.5$ –5 kpc), it is a reasonable first approximation to treat space-time in the loop neighbourhood as being asymptotically Minkowskian. In general a loop will have a peculiar velocity v_p , and also support rapid oscillations⁶, leading to a complicated and time-dependent metric. In the rest frame of the loop centre of mass, the time-dependent terms are essentially gravitational radiation, so that, smoothed over times $t \approx R/c$, and in the region $r \gg R$, the metric will be approximately Schwarzschild. For these reasons I shall take the Schwarzschild metric for $r \gg R$, and 'cut out' the sphere $r < R$ which contains the string loop. The string mass is taken to be $M = \beta \mu R$, with $2\pi \leq \beta \leq 9$, in accordance with estimates based on numerical simulations⁸.

Assuming azimuthal symmetry in the loop rest frame, the nearly homogeneous and isotropic cosmic fluid moves with velocity field $v(r, \theta)$ in the Schwarzschild potential of the loop, with the asymptotic boundary condition that $v \rightarrow -v_p \hat{z}$ as $r \rightarrow \infty$. In any more realistic treatment in which the loop is embedded in an expanding universe, this boundary condition remains essentially unchanged, except that $r \rightarrow \infty$ is replaced by the horizon scale. This can be seen by Lorentz-boosting from the co-moving frame to the rest frame of the loop. I shall further assume that the fluid is perfect, and obeys an equation of state which is implicitly expressible as $p = p(\rho)$, where p and ρ are the fluid pressure and mass-energy density as measured in the fluid rest frame (the co-moving frame). As we are interested here only in small adiabatic perturbations, the sound speed c_s , defined by $c_s^2 = dp/d\rho$, may be regarded as approximately spatially constant.

We seek approximately steady-state solutions to the covariant conservation equation

$$T^{\mu\nu}_{;\nu} = 0 \quad (1)$$

where $T^{\mu\nu}$ is the stress-energy tensor, and the indices μ and ν range from 0 to 3. Defining $\varepsilon = \gamma^2(p + \rho)$, where $\gamma^2 = (1 - v_j^2)^{-1}$ and j ranges from 1 to 3, equation (1) reduces to the familiar equations of steady-state hydrodynamics: continuity,

$$\nabla \cdot (\varepsilon \mathbf{v}) = 0 \quad (2)$$

and the Navier-Stokes equation,

$$\nabla \cdot (\varepsilon \mathbf{v} \mathbf{v}) + \nabla p + \varepsilon \nabla (\ln \sqrt{-g_{00}}) = 0 \quad (3)$$

where $-g_{00} = 1 - 2\Phi_N = 1 - 2MG/r$ in the Schwarzschild metric, and Φ_N is the 'Newtonian' gravitational potential. In fact, equations (2) and (3) are valid in any static and diagonal metric. As $MG/R = \beta \mu G \ll 1$, perturbations in the velocity field are small and hence $\gamma^2 \approx 1/(1 - v_p^2) = \text{constant}$. One can therefore write $p = p(\varepsilon)$, and $dp/d\varepsilon = \tilde{c}_s^2$, where $\tilde{c}_s$ is the effective sound speed. It is reasonable to assume also that the velocity field remains curl-free. In these conditions the Navier-Stokes equation can be integrated to yield the Bernoulli equation:

$$1/2(v^2 - v_p^2) + \tilde{c}_s^2 \ln(\varepsilon/\varepsilon_0) + \ln \sqrt{-g_{00}} = 0 \quad (4)$$

Eliminating ε from the continuity equation by substitution from the Bernoulli equation now gives

$$\nabla \cdot \{v \exp [-(v^2 - v_p^2 + 2 \ln \sqrt{-g_{00}})/2\tilde{c}_s^2]\} = 0 \quad (5)$$

It is sufficient for the present purposes to take a one-dimensional solution as illustrative of the more general solutions of equation (5). To second order in $(v - v_p)/v_p \ll 1$, and expanding $\ln \sqrt{-g_{00}} = \Phi_N$, this is

$$v = -v_p \exp \{ -\Phi_N / (\tilde{c}_s^2 - v_p^2) \} \quad (6a)$$

and

$$\varepsilon = \varepsilon_0 \exp \{ \Phi_N / (\tilde{c}_s^2 - v_p^2) \} \quad (6b)$$

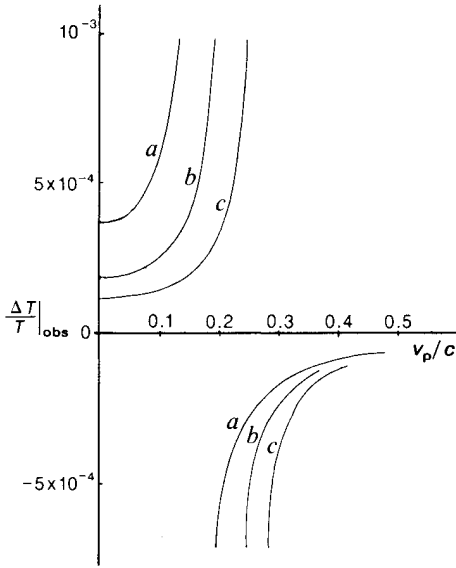


Fig. 1 The CBR microanisotropy due to moving loops of cosmic string, as a function of the loop peculiar velocity, for three different cosmological models parameterized by the baryonic density parameter Ω_b , and with $\mu G/c^2 = 2 \times 10^{-6}$. *a*, $\Omega_b = 1$; *b*, $\Omega_b = 0.5$; *c*, $\Omega_b = 0.3$. The curves shown are discontinuous at $v_p = \tilde{c}_s$, whereas for 'real' loops the effect will be truncated at a finite value determined by the ratio of the sound-speed crossing time to the timescale over which the sound speed changes at recombination.

As there is no assumption of the smallness of $v_p/\tilde{c}_s$ in the derivation of equations (6), it is apparent that: (1) as $v_p \rightarrow 0$, equations (6) reduce to the equation of hydrostatic equilibrium; (2) for $v_p < \tilde{c}_s$, a rather large enhancement of $\varepsilon/\varepsilon_0$ over the hydrostatic case may be expected; and (3) for $v_p > \tilde{c}_s$ the loop actually produces a decrease in the local energy density.

Before recombination the cosmic fluid pressure is dominated by the radiation pressure, so that $\varepsilon = \gamma^2 (\rho_b + 4\rho_\gamma/3)$, and $\tilde{c}_s^2 \approx \{4\gamma^2(1 + \frac{3}{2}d\rho_b/d\rho_\gamma)\}^{-1}$ (in the adiabatic case), where ρ_b is the density due to non-relativistic interacting material, including baryons, and ρ_γ represents the radiation energy density. Particle species which couple only weakly to the baryon-photon fluid, so that their mean free path exceeds the loop radius, should be excluded from the equation of state. Such species do not influence the gas dynamics or the sound speed directly, although they do of course influence the universal expansion rate. 'Cold dark matter' (CDM) components will also tend to cluster around the loop, causing a slight increase in the depth of the potential.

Defining z_{eq} as the redshift at which $\rho_b(z) = \rho_\gamma(z)$, we have

$$\tilde{c}_s^2(z_{eq}) \approx \{4\gamma^2(1 + \frac{9}{16}z_{eq}/z)\}^{-1} \quad (7)$$

In the following numerical estimates, the Hubble constant is taken as $H_0 = 75 h_0 \text{ km s}^{-1} \text{ Mpc}^{-1}$, so that $z_{eq} \approx 2.2 \times 10^4 \Omega_b h_0^2$, where $\Omega_b = \rho_b(0)/\rho_{crit}$ is the baryonic density parameter. Taking the epoch of recombination to occur at $z = 1,500$, the effective sound speed at recombination becomes

$$\tilde{c}_s(z_{rec}) \approx \{4\gamma^2(1 + 8.25\Omega_b h_0^2)\}^{-1/2} \quad (8)$$

Under the foregoing assumptions, equation (6b) implies a local CBR temperature distortion in the hydrostatic limit of

$$\frac{\Delta T}{T}\bigg|_{0,rec} = \frac{1}{3} \frac{\Delta \varepsilon}{\varepsilon}\bigg|_{v_p=0} \approx 1.2 \times 10^{-5} (1 + 8.25\Omega_b h_0^2) \mu_6$$

for $\beta \approx 9$ and $G\mu = 10^{-6} \mu_6$. For a loop with a peculiar velocity $v_{p,rec}$ at recombination, the above result is multiplied by a factor of $(1 - v_{p,rec}^2/\tilde{c}_s^2)^{-1}$. The radiation suffers a gravitational redshift as it climbs out of the loop potential, $\Delta T/T|_{gr} \approx -\Phi_N =$

Table 1 Areal density of string loops for different cosmological models

Ω_0	Ω_m	Angular size of horizon at t_{rec} (deg.)	Areal density of string loops (deg.) ⁻²
1	1	1.15	6.6
1	0.1	2	1.3
0.3	0.3	0.5	27
0.1	0.1	0.22	127

$-10^{-6} \beta \mu_6$, so the final observable temperature distortion is

$$\frac{\Delta T}{T}\bigg|_{v,obs} \approx \left\{ 1.2 \left(\frac{1 + 8.25\Omega_b h_0^2}{1 - v_p^2/\tilde{c}_s^2} \right) - 0.9 \right\} \times 10^{-5} \mu_6$$

as shown for several cosmological models in Fig. 1.

The observable angular distribution and magnitude of CBR distortions caused by loops depends on the number density and velocity distribution of the putative loop population, as well as on the effective sound speed just before recombination, $\tilde{c}_{s,rec}$. Loops are expected to form with peculiar velocities $v_{p,t}$ comparable to the speed of light^{9,10}, which are redshifted so that at a later time $t > t_i$, $v_p(t) = v_{p,t}z(t)/z_i$, where z_i is the redshift of formation. This implies that loops which are initially supersonic will become subsonic before recombination, so long as $z_i/z_{rec} > v_{p,t}/\tilde{c}_{s,rec}$. During recombination the sound speed drops rapidly to $\sim 100 \text{ km s}^{-1}$ over a period $\Delta t_{rec} \approx 5.8 \times 10^4 \text{ yr}$ ($\Delta z_{rec} \approx 400$)¹¹, so that essentially all loops will again become supersonic during recombination. The steady-state analysis leading to equations (6) remains approximately valid so long as the timescale for dynamical equilibration, which is roughly the sound-speed crossing time $t_{sc} \approx R/c_s$, is shorter than the timescale over which the sound speed changes significantly; that is, so long as $(\dot{c}_s/c_s)^{-1} > R/c_s$, or $R/c_{s,rec} < \Delta t_{rec}$. The resulting overdensity $\Delta\varepsilon/\varepsilon_0$ ceases to produce a corresponding CBR distortion when the photon mean free path γ_{mf} becomes comparable to the loop radius. Just before recombination this is $\gamma_{mf} \approx 20(\Omega_b h_0^2)^{-1} \text{ pc}$, and varies approximately as the ionization fraction during recombination. For these reasons one expects loops of radius $R \leq$ a few kpc and with $v_{p,rec} \leq \tilde{c}_s$ to produce the largest distortions in the CBR.

The angular size of loops seen at recombination is $\sim 0.5\Omega_0 h_0 \text{ arc min kpc}^{-1}$, and the appropriate number density for loops with radii in the range R to $R + \Delta R$ is^{6,12} $n(R) = \nu R^{-2} t_{rec}^{-2}$ at recombination. Numerical estimates⁶ put $\nu \approx 10^{-2}$, so taking a range of radii $0.5 < R < 5 \text{ kpc}$, one expects ~ 50 – 70 loops in this range to be within the horizon volume at recombination, and $\sim 10\%$ of these to be contained within the surface of last scattering. The apparent angular size of the horizon at recombination is a function of both Ω_0 and $\Omega_m = \Omega_b + \Omega_{CDM}$, where Ω_{CDM} represents non-interacting, non-relativistic material. The apparent areal density of loops varies as Ω_0^{-2} for a fixed value of Ω_{CDM} , and estimates for several cosmological models are presented in Table 1.

It is clear that a significant fraction of kiloparsec-sized loops might be expected to produce CBR anisotropies of order $\Delta T/T \approx 10^{-3}$ – 10^{-4} in the denser cosmologies, decreasing to $\Delta T/T \geq 4 \times 10^{-5}$ for $\Omega_b h_0^2 \leq 0.1$. Although the magnitude of the anisotropies is well above the best current observational limits on arc-minute scales, the areal density expected in dense cosmologies ($\Omega_0 \geq 0.2$) is so low that detection by small-scale anisotropy studies^{4,5} with small areal coverage is extremely unlikely. Present CBR microanisotropy limits on arc-minute scales are typically based on the observation of $< 10 \text{ (arc min)}^2$ of sky, so that the probability of seeing a loop of angular size comparable to the beam area is $\sim 10^{-2}$ – 10^{-3} . Only for $\Omega_0 < 0.1$ is the areal density of the loops likely to be large enough that they should probably have been detected by now.

Finally, loops with radii of a few hundred parsecs and peculiar velocities close to the sound speed at recombination are capable

of producing rather dramatic density enhancements of order $\Delta\varepsilon/\varepsilon_0 \approx 1$. In such cases the self-gravity of the overdensity itself, which I have so far neglected, will lead to rapid nonlinear collapse and could precipitate black-hole formation on mass scales of $M_{\text{bh}} \approx \rho_{\text{b,rec}} R_{\text{loop}}^3 \approx 5 \times 10^{11} \Omega_{\text{b}} (R/1 \text{ kpc})^3 M_{\odot}$. For the range of loop radii of $100 < R < 300 \text{ pc}$, one expects between

180 to 250 loops per horizon volume at recombination. If only 1% of these satisfied the above criteria one would expect at least one such very massive black hole per horizon volume, which translates to roughly one per 2 Mpc cube now.

This work was supported by an SERC postdoctoral research fellowship.

Received 20 February; accepted 4 June 1986.

1. Kibble, T. W. B. *J. Phys. A* **9**, 1387–1398 (1976).
2. Vilenkin, A. *Phys. Rev. D* **24**, 2082–2089 (1981).
3. Zeldovich, Ya. B. *Mon. Not. R. astr. Soc.* **192**, 663–667 (1980).
4. Uson, J. M. & Wilkinson, D. T. *Nature* **312**, 427–429 (1984).
5. Lasenby, A. N. & Davies, R. D. *Mon. Not. R. astr. Soc.* **203**, 1137–1170 (1983).

6. Brandenburger, R. H. & Turok, N. *Inst. theor. Phys. Preprint No. NSF-ITP-85-88* (1985).
7. Kaiser, N. & Stebbins, A. *Nature* **310**, 391–393 (1984).
8. Albrecht, A. & Turok, N. *Phys. Rev. Lett.* **54**, 1868–1871 (1985).
9. Kibble, T. W. B. & Turok, N. *Phys. Lett.* **116B**, 141–143 (1982).
10. Vilenkin, A. & Shafi, Q. *Phys. Rev. Lett.* **51**, 1716–1719 (1983).
11. Peebles, P. J. E. *Physical Cosmology* (Princeton University Press, 1971).
12. Vilenkin, A. *Phys. Rep.* **121**, 263–315 (1985).

Discovery of organic grains in comet Halley

D. T. Wickramasinghe* & D. A. Allen†

* Department of Mathematics, Australian National University, Canberra, ACT 2600, Australia

† Anglo-Australian Observatory, Epping, NSW 2121, Australia

Comets are believed to have formed at the same time as the other bodies in the Solar System, but at a sufficient distance from the proto-Sun that their composition was unaffected by fractionation. Thus, they may preserve intact the chemical composition of the original solar nebula, and perhaps of much of the Galaxy. Previously we demonstrated^{1,2} that one component of the interstellar dust grains has an organic composition, by detecting a broad absorption band at a central wavelength of $3.4 \mu\text{m}$ along the line of sight to IRS7, a luminous star near the galactic centre. We have performed the same observations of comet Halley, and now report the detection of a similar emission band in the comet's spectrum. This evidence of organic grains in comets indicates the possible presence therein of a component of similar composition to that found in micrometeorites. Although no definite identification is possible we note that a feature near $3.4 \mu\text{m}$ is predicted by various models of interstellar dust^{3–7}.

The observations were made with the 3.9-m Anglo-Australian Telescope (AAT), using the infrared photometer-spectrometer in its spectroscopic mode. We adopted the observing technique used for IRS7, namely, a series of 32 spectral scans of the object interspersed with identical scans of a nearby bright star. Our data were thus almost immune to changes in the atmospheric transmission. Each series of scans occupied $\sim 10 \text{ min}$ of observing time, and four such series were observed each night.

Because of the size of the comet, sky reference had to be taken some considerable distance away. We chose a beam separation of 100 arc s and drove the telescope between comet and sky after every scan through the spectrum, approximately every 10 s. This technique relies on the stability of the sky, which is not always adequate. If the sky has varied, then a proportion of its spectrum will be added to or deducted from that of the comet. Such instances are readily recognized by their distinct mean flux. In reducing the data, we have rejected individual spectra for which variation of the sky clearly caused appreciable degradation. Usually none were rejected, but in the worst case (late on the third night), we rejected half of the data.

Data were taken on three successive nights, March 30, 31 and 1 April UT. On each occasion we centred our entrance aperture on the position of greatest infrared flux, which corresponded to the visible nucleus. On the first night the comet appeared quite faint through a 5-arc-s circular aperture. For the other two nights, therefore, we employed a rectangular aperture measuring $10 \times 5 \text{ arc s}$. The long axis lay in declination. The spectral resolution, nominally $\lambda/\Delta\lambda \approx 100$, was not degraded below ~ 90 for these apertures unless the distribution of emitting dust differed greatly

from that of the reflecting dust providing the visible object.

On the second night the comet was much brighter, and on the third night it had faded to an intermediate level. On the second night the ratio of fluxes through the two apertures was 1.6. This factor was applied to the first night's data to allow comparison of the three data sets, although we have no guarantee that that factor was appropriate to the first night.

A nearby bright star, $\theta \text{ CrA}$, was used as a secondary standard for the principal wavelength range observed, $2.9\text{--}4.0 \mu\text{m}$. Observations of the comet and the star were alternated every 10–15 min. $\theta \text{ CrA}$ was in turn calibrated against one of the standard spectrophotometric stars in use at the AAT⁸, BS6748. Data on the comet were also secured in the shorter-wavelength atmospheric windows $1.4\text{--}1.8$ and $2.0\text{--}2.5 \mu\text{m}$, and were directly calibrated using BS6748.

Figure 1 shows the data taken on the second night. Within the two shorter-wavelength atmospheric windows, reflected solar radiation dominates, and the spectrum is featureless with a gradient close to that of the Sun. Longward of $3 \mu\text{m}$, thermal emission from the dust is seen. A satisfactory fit to this wavelength region is afforded by the sum of the extrapolated solar component and a 350 K ($\pm 10 \text{ K}$) black body. This temperature is higher than the equilibrium value at the solar distance of 1.17 AU , as expected for the smaller grains, which are not effective radiators at longer infrared wavelengths, and which consequently rise to temperatures above the black-body equilibrium. Figure 1 shows the fitted continuum (dashed line); above this rises a broad emission band centred at $3.4 \mu\text{m}$. In this and the other data shown here, we have omitted the data points in the narrow telluric methane absorption bands, since these cannot, in general, be adequately corrected for when an extended object is observed.

Figure 2 compares the $3.4\text{-}\mu\text{m}$ data for the three nights. All three have been plotted at the same scale, and the different levels are due entirely to variation of the comet. The continuum plotted for the 31 March data is that shown in Fig. 1; for the other dates, the same continuum has been adjusted in intensity to fit. The intensity of the $3.4\text{-}\mu\text{m}$ band relative to the continuum is not constant; hence, we infer that there are two unrelated grain components in the comet's ejecta, which produce the continuum and the emission band separately.

We subtracted the adopted continuum from each night's data. The resultant emission feature was found to be of identical intensity on the second and third nights, and about a factor of 4.5 weaker on the first night. The profile is seen in Fig. 2 to comprise two broad peaks centred at ~ 3.38 and $3.51 \mu\text{m}$. Following our initial announcement⁹, a higher-resolution spectrum by Geballe¹⁰ was described as having peaks at 3.36 and $3.52 \mu\text{m}$. Another high-resolution observation ($\lambda/\Delta\lambda \approx 400$) was made by D.A.A., P. F. Roche and B. Bailey (unpublished data) on 24 April using an infrared grating spectrometer on the AAT. The comet was then faint and the observation brief, but the two peaks were again seen.

At the resolution of the present data, the emission seems most closely to match the absorption in the line of sight to $\rho \text{ Oph} 29$ (ref. 11), a source in a dark cloud, but there are general

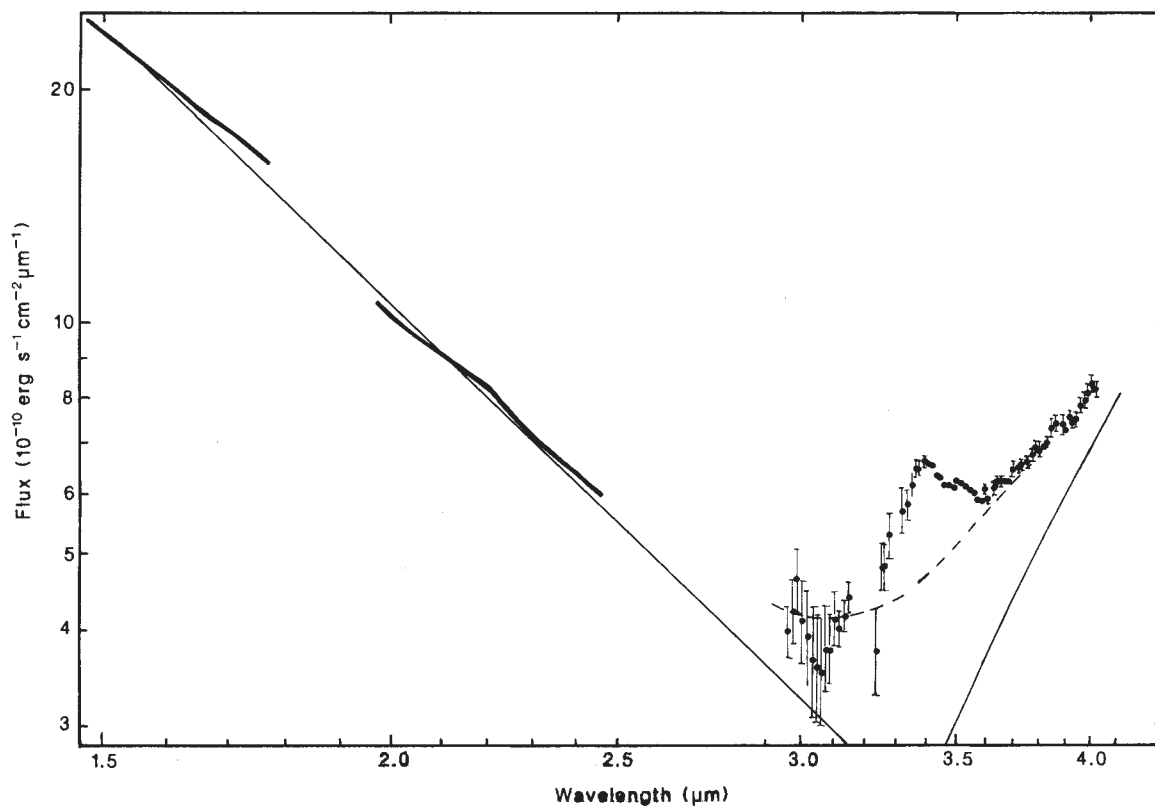


Fig. 1 The infrared spectrum of comet Halley on 31 March 1986. The data in the two shorter-wavelength windows are featureless. Dashed line, the adopted continuum derived from a sum of a 350 K black body and the scattered sunlight.

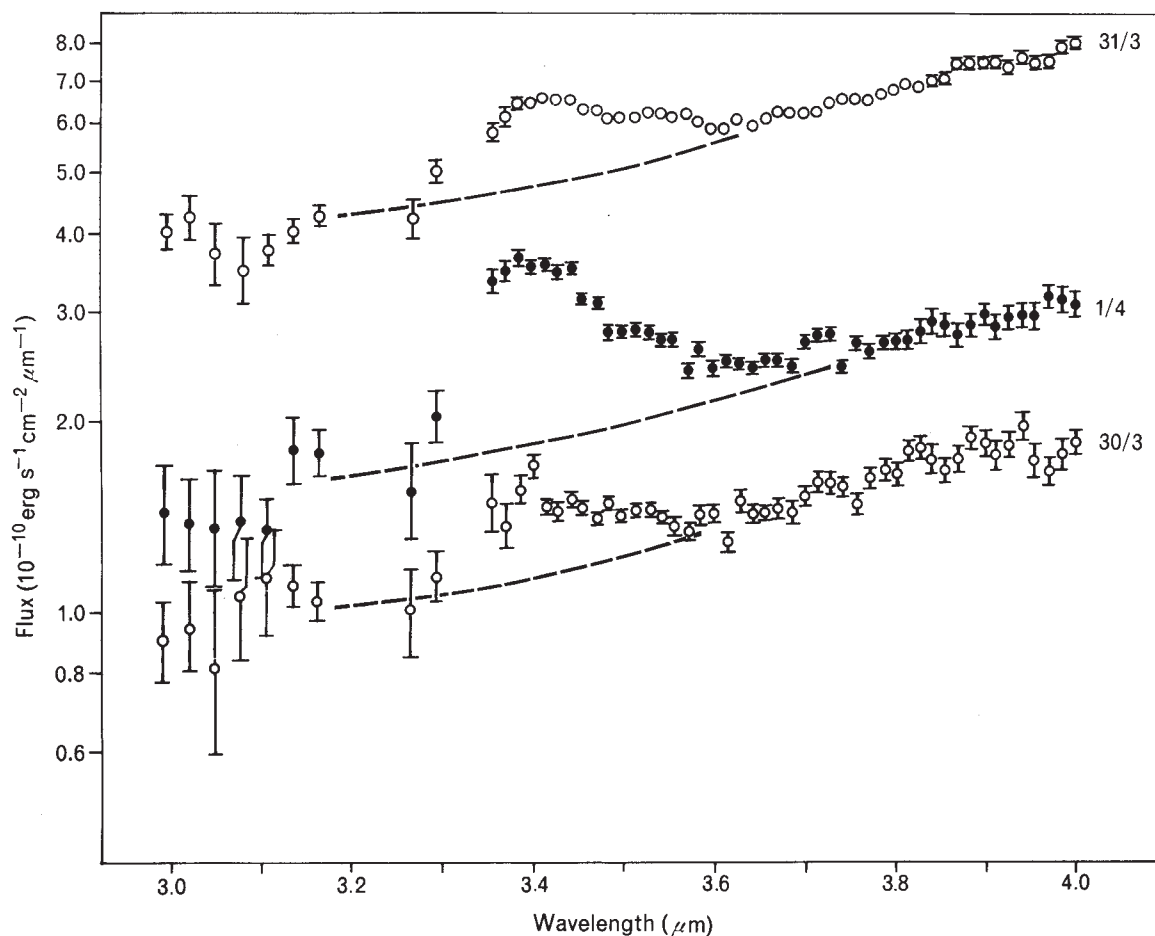


Fig. 2 Spectra at 3.0–4.0 μm for the three dates of observation. Apart from a correction for the different aperture used on 30 March, all three data sets are at the same flux scale. Shortward of 3.35 μm , the data have been lightly smoothed.

similarities to the absorption in most sight lines tested. Emission in the band 3.25–3.55 μm is common in galactic nebulae and stars¹², but is nowhere similar to the emission in comet Halley.

Whether the material in comet Halley is identical to that in any of the sight lines through the Galaxy, or merely similar, the same argument for an organic origin pertains. The central wavelength is in the range expected for C–H bond vibrational oscillations, and its structureless character indicates a likely origin in the solid phase. We cannot, however, exclude a contribution from single molecules of moderately high molecular weight, in which the fine structure would also be blended.

The suggestion that organic material might be found in comets is not new^{3,13,14}. Impetus has been given to the idea by the discovery that some comets (such as Crommelin¹⁵) have very dark nuclei, a fact best explained by a tarry organic residue on the surface. A very low albedo for comet Halley was found by the Vega and Giotto spacecraft^{16,17}. Giotto also found a large fraction of the smaller grains ($\leq 1 \mu\text{m}$) to be composed mainly of H, C, N and O, typical of organic materials¹⁸. Infrared spectroscopy of sufficient quality to reveal the 3.4- μm emission band has not previously been published for any comet, although there is a hint of an emission feature in the published spectrum of the nucleus of comet IRAS-Araki-Alcock¹⁹.

The presence of the emission band is probably explained by particle size effects. As noted above, smaller particles adopt higher temperatures, thus radiating more strongly. It is the smaller particles which give rise to the emission band, for in large grains the spectral signature is diluted, eventually to the black-body continuum. Indeed, it is possible that the two components we see, producing continuum and band, are chemically the same, and distinguished by size alone.

Although a definite identification of the material responsible for the 3.4- μm band cannot be made from the present data, the observations can be used to test and place constraints on various models. We note that hydrogenated amorphous carbon and polycyclic aromatic hydrocarbons have been proposed to explain spectral features in this waveband⁷. Other possible candidate materials are the non-volatile organic residues that have been produced in Urey-Miller-type experiments carried out by Khare and Sagan⁴, Moore and Donn⁵ and Greenberg⁶. We note that the Hoyle-Wickramasinghe³ biological grain model for cometary dust makes a specific prediction for the 3.4- μm feature which can potentially be tested with the present data. It may be significant that the structured organic component of the Murchison meteorite has a 3.4- μm feature which is similar to that seen in Halley²⁰. The Murchison meteorite contains complex organic molecules, including various amino acids²¹; the present observations may indicate that similar material is found in the organic dust of comets.

The organic material in comets, and throughout our Galaxy, may soon be identified. Microscopic grains, known as Brownlee particles, have been collected in the stratosphere, and techniques for their analysis are being developed. In a study of the infrared transmission spectrum of some of these particles, Sandford and Walker²² found several which show a feature at 3.4 μm , even though they deliberately sought to study chondritic specimens. They could not eliminate the possibility of contamination from the silicone oil used to capture the grains. More careful experiments directed at the darker materials adhering to some chondritic grains may ultimately lead to the identification we seek.

We thank Andrew Longmore for making available telescopic time on 1 April, and Martha Hanner for helpful discussions. D.T.W. acknowledges support from the Australian Research Grant Scheme and the Faculties Research Fund.

6. Greenberg, J. M. *Molecules in the Galactic Environment* (eds Gordon, M. A. & Snyder, L. E.) 94–118 (Wiley, New York, 1973).
7. Duley, W. W. & Williams, D. A. *Mon. Not. R. astr. Soc.* **205**, 67P–70P (1983).
8. Allen, D. A. & Cragg, T. A. *Mon. Not. R. astr. Soc.* **203**, 777–783 (1983).
9. Wickramasinghe, D. T. & Allen, D. A. *IAU Circ. No.* 4205 (1986).
10. Geballe, T. *IAU Circ. No.* 4212 (1986).
11. Wickramasinghe, D. T. & Allen, D. A. *Astrophys. Space Sci.* **97**, 369–378 (1983).
12. Allen, D. A., Baines, D. N. T., Blades, J. C. & Whittet, D. C. B. *Mon. Not. R. astr. Soc.* **199**, 1017–1024 (1982).
13. Strazzulla, G. *Icarus* **61**, 48–53 (1985).
14. Greenberg, J. M. *Comets* (ed. Wilkening, L. L.) 131–163 (University of Arizona Press, 1982).
15. Hanner, M. S., Knacke, R. F., Sekanina, Z. & Tokunaga, A. T. *Astr. Astrophys.* **152**, 177–181 (1985).
16. Sagdeev, R. Z. *et al. Nature* **321**, 262–266 (1986).
17. Keller, H. V. *et al. Nature* **321**, 320–326 (1986).
18. Kissel, J. *et al. Nature* **321**, 280–282 (1986).
19. Hanner, M. S. *et al. Icarus* **62**, 97–109 (1985).
20. Hoyle, F. & Wickramasinghe, N. C. *From Grains to Bacteria*, 182 (University College Cardiff Press, 1984).
21. Engel, M. H. & Nagy, B. *Nature* **296**, 837–840 (1983).
22. Sandford, S. A. & Walker, R. M. *Astrophys. J.* **291**, 838–851 (1985).

Mobilization of cryogenic ice in outer Solar System satellites

David J. Stevenson* & Jonathan I. Lunine†

* Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

† Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

Voyager images¹ of the uranian satellites show a diversity of geological features, including clear evidence for the 'softening' and mobilization of ice on Miranda and Ariel. Some of these features are similar to those seen on jovian and saturnian satellites, where the mobile material is believed to be water or a water-ammonia mixture. However, the extremely low temperatures and probable unavailability of large energy sources within the uranian satellites lead us to consider flow mechanisms that operate at very low temperature ($T \leq 100 \text{ K}$). We propose here a form of pressure-solution creep, in which very fine-grained water ice or clathrate hydrate is mobilized by a small amount of intergranular cryogenic fluid (CH_4 , CO or N_2). Viscosities as low as 10^{12} P are possible for a limited time, sufficient to allow flooding of rift valleys and perhaps even substantial lateral flows (glaciers).

Preliminary results¹ indicate average densities for the uranian satellites of $\sim 1.6 \text{ g cm}^{-3}$, significantly larger than the values typical for small saturnian satellites ($\sim 1.3 \text{ g cm}^{-3}$). This may be explained² by a uranian satellite formation environment in which a substantial fraction of the carbon is oxidized (primarily CO) rather than reduced (CH_4), thereby decreasing the amount of H_2O , the primary ice condensate. The satellite compositions are then a mixture of clathrate³ ($\sim \text{X} \cdot 6\text{H}_2\text{O}$; $\text{X} = \text{CO}$, N_2 , CH_4) and rock. If the formation conditions were sufficiently cold, then some of the more volatile ices (CO , N_2 and CH_4) could have condensed directly, but these are so volatile that the gravitational energy of satellite formation may have prevented efficient retention², except in very small bodies such as Miranda. To the extent that 'fragile' volatiles may aid mobilization and tectonics, it may actually help to have a small body (the antithesis of the usual expectation that large bodies are the most likely to have internal activity). The biggest puzzle with the uranian satellites is the nature and magnitude of the energy source(s) responsible for resurfacing. We do not attempt to resolve this puzzle here, but adopt the working hypothesis that only a small energy source was available (sufficient to heat near-surface ice by a few tens of degrees from the ambient sub-surface average of $T \geq 60 \text{ K}$) and seek to identify a mobilization mechanism which works at $T \leq 100 \text{ K}$. The presence of sharp, unrelaxed bowl craters in terrains immediately adjacent to resurfaced areas means that this mechanism must be transient and selective.

At $T \leq 100 \text{ K}$, pure water ice or clathrate and probably any ammonia hydrate are essentially 'rock-like' in their rheological

Received 14 May; accepted 1 July 1986.

1. Wickramasinghe, D. T. & Allen, D. A. *Nature* **287**, 518–519 (1980).
2. Allen, D. A. & Wickramasinghe, D. T. *Nature* **294**, 239–240 (1981).
3. Hoyle, F. & Wickramasinghe, N. C. *Astrophys. Space Sci.* **66**, 77–90 (1979).
4. Khare, B. N. & Sagan, C. *Icarus* **20**, 311–321 (1973).
5. Moore, M. H. & Donn, B. *Astrophys. J.* **257**, L47–L50 (1982).

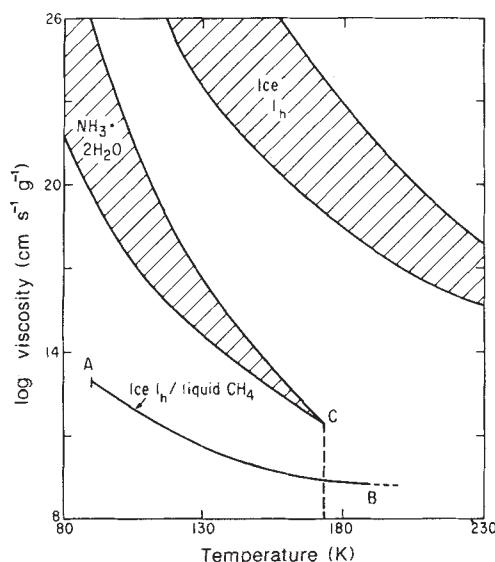


Fig. 1 Logarithm of viscosity versus temperature of three candidate materials for outer Solar System satellites. The uppermost region is the viscosity of pure hexagonal water ice, assuming a newtonian rheology and two choices of activation energy and viscosity at the melting point²³. The middle region is an ammonia-water solid with composition equal to that of the water-rich eutectic; the viscosity at the freezing point C is assumed to be 10 orders of magnitude higher than that of the eutectic liquid²², and the range of activation energies is taken to be equal to that for water ice I_h (ref. 23). Lower curve is the viscosity of water ice I_h with liquid methane in the spaces between grains, calculated using the pressure-solution creep mechanism described in the text. Points A and B, liquid-methane freezing and critical points, respectively.

behaviour, meaning that the Maxwell time, defined as the ratio of viscosity to shear modulus, exceeds the age of the Solar System. This is because the activation energy ΔE of any plausible creep mechanism is large compared with the thermal energy, kT , and the viscosity $\eta \propto \exp(\Delta E/kT)$. A similar argument leads to the conclusion that common terrestrial mantle minerals (such as olivine) exhibit negligible creep at $T \leq 1,000$ K. However, creep in terrestrial materials can be enhanced enormously by the presence of an appropriate intergranular fluid⁴⁻⁷. In this process, known as pressure-solution creep, the fluid is the conduit for the redistribution of the constituent molecules of the grains. In the system quartz (SiO_2)-water, for example, SiO_2 is preferentially dissolved at intergrain boundaries (a high-stress environment) and deposited in regions of low stress, and the creep rate is a function of the pressure dependence of the solubility of quartz in water. In our application, the quartz is replaced by small grains of water ice or clathrate and the liquid water is replaced by liquid cryogen, CH_4 , N_2 or CO .

An approximate expression for the resulting creep is^{5,8}

$$\eta \approx \frac{h^3 \rho_s \sigma_s}{24 \delta \rho_l D} \quad (1)$$

where h is the grain diameter, ρ_s is the density of the solid phase, ρ_l is the density of the liquid, δ is the width of the grain boundary, and D is the diffusion coefficient for the solute (molecules of the grain) in the liquid. The stress parameter σ_s is defined by

$$C(\sigma, T) = C_0 + \sigma/\sigma_s \quad (2)$$

where C is the solubility of the grain material in the liquid (expressed as a mole fraction) and σ is the local deviatoric stress. If we assume ideal solution theory, then (following ref. 5) we have $C/C_0 = \exp(V\sigma/RT)$, where V is the molar volume in the solid phase and R is the gas constant. Provided $\sigma \ll RT/V$ (~ 500 bar for water ice at 100 K), we recover equation (2) with

Table 1 Comparison of terrestrial and outer planet satellite conditions

Solid matrix	Terrestrial example Quartz (SiO_2)	Outer Solar System example Water ice or clathrate
Intergranular fluid	H_2O	CH_4
Grain size (cm)	0.2	10^{-4}
σ_s (dyn cm^{-2})*	3×10^9	$\sim 10^{14}$
D ($\text{cm}^2 \text{s}^{-1}$)	2×10^{-4}	2×10^{-5}
Viscosity (P)†	$\sim 10^{17}$	$\sim 10^{13}$
	(773 K)	(90 K)

* Equation (2).

† Equation (1).

$\sigma_s \approx RT/VC_0$. We are interested in flows driven by $\sigma < 1$ bar. The low-pressure solubility C_0 has been measured by Rebiai *et al.*⁹ for three solvents (CH_4 , N_2 , O_2) but at only a few temperatures. For example, the solubility of H_2O in N_2 at 77.4 K is 1×10^{-5} , about the same as the solubility of quartz in water at room temperature¹⁰. We find that we can explain all the cryogenic data by a model based on ideal solution theory, in which it is assumed that the solid phase is clathrate (CL) and the standard state for water is a fictive non-hydrogen-bonded liquid (non-H liq.). The equality of chemical potentials can then be expressed as

$$f(\text{H}_2\text{O}; \text{CL}) = C_0 f(\text{H}_2\text{O}; \text{non-H liq.}) \quad (3)$$

where f is the fugacity of water in clathrate (left-hand side) and fictive fluid (right-hand side), respectively. The adjustable parameters are the enthalpy of conversion, ΔH , of ice I to the fictive standard state and the difference in specific heats at constant pressure, Δc_p , between the two states. Fitting ΔH to the Rebiai *et al.*⁹ data for various plausible choices of Δc_p (ref. 11) yields similar temperature dependences for $C_0(T)$. We choose for concreteness $\Delta c_p = 14 - 0.033T$ cal $\text{mol}^{-1} \text{K}^{-1}$, which yields $\Delta H = 3,800$ cal mol^{-1} . This model enables us to calculate the temperature dependence of C_0 for N_2 and CH_4 solvents. (From general physical chemical principles¹², we can expect that the behaviour for CO will be similar to that for N_2 .)

The interdiffusion coefficient D is $\sim 2 \times 10^{-5} \text{ cm}^2 \text{s}^{-1}$, based on the application of the Stokes-Einstein equation to shear viscosity measurements on liquid argon^{13,14}. The grain size h may be very small because the water ice has been unable to metamorphose at the extremely low temperatures obtaining before and after satellite formation. Grain sizes $\leq 1 \mu\text{m}$ are inferred for interstellar material¹⁵, comets¹⁶ and primitive meteoritic material (Brownlee particles)¹⁷, and are assumed here. (The water ice may be amorphous, but each grain will develop a clathrate coating.³) The remaining major uncertainty is the grain-boundary width. We adopt $\delta = 10^{-7} \text{ cm}$, as has been used in terrestrial applications^{5,8}, but this is very uncertain. It cannot be much smaller if the model is to have any validity. In Table 1, we compare the parameters characterizing creep in $\text{H}_2\text{O}(\text{s})$ - CH_4 at 90 K with the possible behaviour of SiO_2 (quartz)- H_2O at 773 K. Figure 1 shows the temperature dependence of viscosity for water ice, $2\text{H}_2\text{O} \cdot \text{NH}_3$ (eutectic mixture) and the medium described here. There are essentially no data for $2\text{H}_2\text{O} \cdot \text{NH}_3$ below 175 K, other than the observations of glass formation¹⁸. The graph reflects the theoretical expectations¹⁹ that the viscosity will rise many (~ 10) orders of magnitude during glass formation and continue to rise at still lower temperatures because of the large energy of activation in hydrogen-bonded materials. The main point of Table 1 and Fig. 1 is that, provided the grain size is small, pressure-solution creep at $T \leq 100$ K can produce viscosities comparable to or less than that appropriate for temperate glaciers on Earth²⁰. Of course, this behaviour is truncated at $T \leq T_m$, the melting point of the cryogen. The examples shown here are for a methane intergranular fluid, but similar results apply for N_2 , CO or a mixture.

There are two potential problems with applying the ideas presented here: how do we create the right environment and how do we maintain this environment for a sufficiently long time? A water-ice matrix with intimately penetrating cryogenic fluid is not the thermodynamically preferred state in situations of interest, because of clathrate formation³. The required environment could be created either by clathrate saturation (more cryogen present than can be incorporated in a compound of the form $\sim X \cdot 6H_2O$) or by a disequilibrium state (potentially short-lived), for example, the invasion of water ice by cryogen flooding from adjacent regions. Dissociation of clathrate can occur by heating episodes²¹, but the released cryogen is in gas form and would not liquefy unless trapped (possible if the overlying ice is impermeable). Clathrate dissociation is a plausible trigger for mobilizing ice because it causes expansion. Fracture and rifting of the overlying ice can then occur, with the mobilized medium flowing in response to the hydrostatic pressure head. Flow would persist until the intergranular fluid either escaped or froze. Escape rates can be estimated from Darcy's law (ref. 8, p. 413), which predicts fluid velocities of $\sim 10^{-10} \text{ cm s}^{-1}$, negligibly small because of the small grain size. We can estimate the maximum lateral flow distance before freezing by evaluating the flow velocity v and the cooling time τ_{cool} :

$$v \leq 0.1 \rho_s g \alpha d^2 / \eta \quad (4)$$

$$\tau_{\text{cool}} \leq d^2 / K \quad (5)$$

where d is the flow thickness, α is the local slope, g is the gravitational acceleration and K is the thermal diffusivity. For the choices $g \approx 10 \text{ cm s}^{-2}$ (appropriate to Miranda), $\alpha = 10^{-2}$, $K = 10^{-2} \text{ cm}^2 \text{ s}^{-1}$, the maximum flow distance $L \approx v \tau_{\text{cool}}$ is given by

$$L \approx 10 \text{ km } (d/300 \text{ m})^4 (10^{12}/\eta) \quad (6)$$

This crude estimate suggests that modest (but rather thick) flows are possible, probably sufficient at least to flood rift valleys.

We have presented an idea, not an explanation. Experimental work is desirable and more work is needed on the nature of the composition and energy sources in these satellites. We believe, however, that the presence of geological activity on small outer Solar System satellites is not so surprising when viewed from the perspective of the physical chemistry of cryogenic materials.

This work was supported by NASA grant NAGW185 and NSF grant AST8206173. Contribution 4323 from the Division of Geological and Planetary Sciences, California Institute of Technology.

Tropical climatic phase lags and Earth's precession cycle

David A. Short* & John G. Mengel†

* Goddard Laboratory for Atmospheres/Code 613, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

† Applied Research Corporation, Landover, Maryland 20785, USA

The Earth's precession cycle causes changes in the seasonal cycle of incident solar radiation, affecting high-latitude and tropical climate in the period range $\sim 19\text{--}24 \text{ kyr}$ (refs 1–3). Atmospheric general circulation models (GCMs) have shown that thermally forced monsoon circulations are enhanced by the strengthened seasonal cycle of solar heating that results when summer solstice occurs near perihelion^{4,5}. However, observations indicate that the tropical climatic response reaches a maximum thousands of years after solstice–perihelion alignment^{2,6}. Because the GCM experiments included slowly evolving boundary conditions (such as ice sheet extent and surface albedo), whose influences on tropical climate are not well understood, reasons for the observed lag are unclear. We have addressed this uncertainty by examining the response to the precession cycle. Our model suggests that some features of the monsoon circulation may reach maximum intensity $\sim 3,000 \text{ yr}$ after solstice–perihelion alignment as the result of a direct thermal response to the astronomical forcing.

Our climate model is a seasonal, two-dimensional, linear energy-balance model (EBM) with diffusive heat transport. The only prescribed zonal asymmetry is the distribution of land, sea ice and ocean through their respective heat capacities. The longest internal timescale in the model is $\sim 5 \text{ yr}$, due to a 75-m ocean mixed-layer depth. The model parameters and a favourable comparison with the Earth's present annual and semi-annual cycles of surface temperature have been compiled elsewhere⁷. Small perturbations in the seasonal cycle of solar forcing caused by orbital variations produce seasonal temperature perturbations in the EBM and in more comprehensive GCMs. The sensitivity to orbital variations of the EBM and GCMs have been shown to be comparable⁸. Because monsoon circulations are primarily forced by the temperature contrast that develops between land and ocean in response to the seasonal cycle of solar heating, we can infer the sensitivity to orbital variations of this component of the climate system. A thorough investiga-

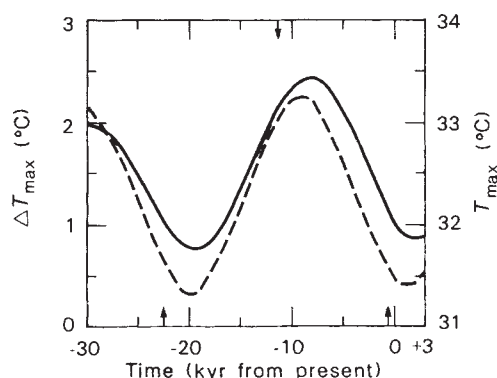


Fig. 1 Time series (dashed line) of the modelled maximum temperature during the year for 15° N , 20° E (north central Africa). Vertical arrows at the top and bottom mark times when summer (top) or winter (bottom) solstice occurred at perihelion. Also shown (solid line) is the maximum temperature difference that develops each summer between 15° N , 20° E and the tropical Atlantic Ocean (0° N , 0° E). Because thermal contrast is an important driving force for tropical atmospheric circulations, these thermal fields may be considered to be indicative of moisture availability to north central Africa during the rainy season.

Received 10 April; accepted 29 May 1986.

- Smith, B. A. *et al. Science* **233**, 43–64 (1986).
- Stevenson, D. J. in *Uranus and Neptune* (ed. Bergstrahl, J. T.) 405–423 (NASA Conf. Publ. No. 2330, 1984).
- Lunine, J. I. & Stevenson, D. J. *Astrophys. J. Suppl. Ser.* **58**, 493–531 (1985).
- Fyfe, W. S., Turner, F. J. & Verhoogen, J. *Mem. geol. Soc. Am.* No. 73 (1958).
- Rutter, E. H. *Phil. Trans. R. Soc. A* **283**, 203–219 (1976).
- Durney, D. W. *Phil. Trans. R. Soc. A* **283**, 229–240 (1976).
- Pharr, G. M. & Ashby, M. F. *Acta metall.* **31**, 129–138 (1983).
- Turcotte, D. L. & Schubert, G. *Geodynamics*, 335 (Wiley, New York, 1982).
- Rebiai, R., Rest, A. J. & Scurlock, R. G. *Nature* **305**, 412–413 (1983).
- Siever, R. *J. Geol.* **70**, 127–150 (1962).
- Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water*, 99–175 (Oxford University Press, 1969).
- Vegard, L. *Z. Phys.* **61**, 185–190 (1930).
- Rudenko, N. W. & Schubikow, L. W. *Phys. Z. SowjUn.* **6**, 470 (1934).
- Longuet-Higgins, H. C. & Pople, J. A. *J. chem. Phys.* **25**, 884 (1956).
- Draine, B. T. in *Protostars and Planets II* (eds Black, D. C. & Matthews, M. A.) 621–640 (University of Arizona Press, Tucson, 1985).
- Ney, E. P. in *Comets* (ed. Wilkening, L. L.) 323–340 (University of Arizona Press, Tucson, 1982).
- Brownlee, D. E. in *Protostars and Planets* (ed. Gehrels, T.) 134–150 (University of Arizona Press, Tucson, 1978).
- Van Kasteren, P. H. G. *Bull. Inst. int. Froid. Annexe 4*, 81–87 (1973).
- Weyl, W. A. in *Rheology: Theory and Applications* Vol. 3 (ed. Eirich, F. R.) 299–340 (Academic, New York, 1960).
- Hobbs, P. V. *Ice Physics* Ch. 4 (Clarendon, Oxford, 1974).
- Stevenson, D. J. *Nature* **298**, 142–144 (1982).
- McKinnon, W. B. & Meadows, M. *Bull. Am. astr. Soc.* **16**, 686 (1984).
- Friedson, A. J. & Stevenson, D. J. *Icarus* **56**, 1–14 (1983).

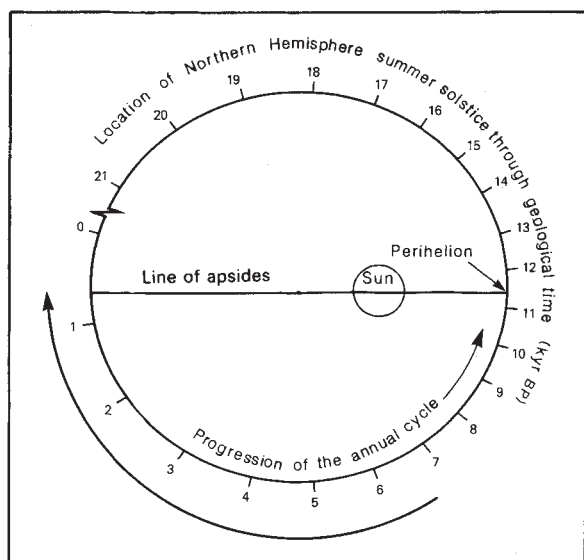


Fig. 2 A chronology of the relationship between summer solstice and perihelion for the most recent precession cycle. Locate geological time on the outer scale, determining the position of the northern summer solstice; then, proceed in the direction of the inner arrow to simulate Earth's annual orbit around the Sun. For example, at 8.5 kyr BP the northern summer solstice occurred ~6 weeks (1/8 of the distance around the dial = 1/8 of a year) before perihelion. For illustrative purposes the eccentricity of this ellipse is 0.30.

tion of the properties of our computationally inexpensive model may shed light on the behaviour of the Earth's climate as inferred by more comprehensive models or as observed in the geological record.

The evolution of the seasonal thermal field was examined in a series of model runs simulating orbital conditions from 30 kyr before present (BP) to 3 kyr after present. During this time interval a complete precession cycle takes place, the eccentricity hovers just below 0.02 and the obliquity reaches a maximum at 10 kyr BP. Figure 1 shows a time series of maximum summer temperature (MST) over north central Africa. Also shown is the maximum temperature contrast that develops each year between north central Africa and the tropical Atlantic Ocean. These thermal fields can be regarded as being indicative of the moisture availability to the area during the rainy season. The peak response occurs between 8 and 9 kyr BP, in remarkable agreement with the evidence of palaeolake levels⁶, and several thousand years after solstice-perihelion alignment. A similar lag was found in minimum temperatures during winter (the season dominated by evaporation), in the sense that would reinforce a response dominated by the precipitation season. The lag was also found to apply to seasonally averaged temperatures centred around MST. Thermal contrast between north Africa and the Indian Ocean, which helps support the low-level south-west flow into India, was also maximum at ~8.5 kyr BP. Although thermal contrast is an over-simplification of monsoon dynamics, the EBM results warrant further study of orbital forcing with dynamical models.

Further experiments showed that MST and land-sea temperature contrast in the tropics were maximum at 8–9 kyr BP, due to effects of the Earth's eccentric orbit and the precession cycle. Perihelion at 8.5 kyr BP occurred ~6 weeks after summer solstice. At 15° N the Sun passes overhead ~6 weeks before and after summer solstice. Roughly speaking, in our model the warmest summers over tropical land occur at a time when perihelion coincides with the second overhead passage of the Sun.

The lag between warmest summers and solstice-perihelion alignment in our linear model can be explained more precisely.

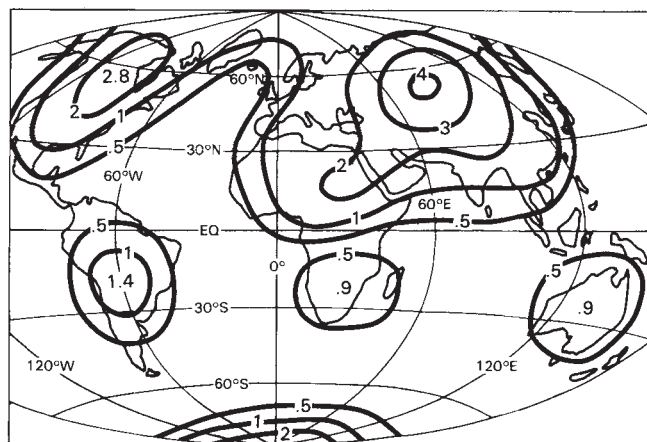


Fig. 3 Temperature difference (°C) between the warmest and coolest summers as the precession cycle moves northern summer solstice from perihelion to aphelion and back again with an eccentricity of 0.02 and an obliquity of 23.25°. For an eccentricity of 0.04 (0.06) the contours should be re-scaled by a factor of 2 (3).

Consider first a model planet with a circular orbit, and a tropical latitude where the amplitudes of the annual and semi-annual cycles in solar forcing are equal. For the annual cycle of solar forcing the thermal response over land lags the forcing by ~30 days. The semi-annual response amplitude is 75% of the annual response and the lag is 24 days. Over the ocean, phase lags are $\frac{1}{4}$ cycle for each frequency (~90 and 45 days), while the response amplitude at the annual cycle is $\frac{1}{30}$ that over land and the semi-annual response is $\frac{1}{60}$. Note that the lag of the thermal response is dependent on local heat capacity and frequency. Maximizing a local thermal response by aligning the phases of the response frequencies requires that the phases of the forcing frequencies be misaligned. However, for a circular orbit the phases of the forcing frequencies are fixed by the timing of the summer solstice. The varying Earth-Sun distance in an eccentric orbit introduces new terms into the seasonal forcing, with phases that depend on the relative positions of summer solstice and perihelion. A maximum temperature response can now be realized by a misalignment of the phases of the forcing frequencies. In the context of our model experiments, the misalignment is revealed as an offset between summer solstice and perihelion. The optimal offset will depend not only on heat capacity but also on latitude, because the annual cycle of forcing dominates outside the tropics. An offset of 50° occurred at 8.5 kyr BP, placing perihelion ~6 weeks after summer solstice. Our model shows that this orbital configuration, occurring ~3 kyr after solstice-perihelion alignment, produces the warmest summers and greatest land-sea temperature contrast in the tropics. Figure 2 shows the geometrical relationships between summer solstice, the annual cycle and perihelion for the most recent precession cycle.

To examine geographical patterns of the influence of the precession cycle on surface temperature, a series of model runs for one precession cycle were made with a fixed obliquity (23.25°) and eccentricity (0.02). At each geographical location the temperature difference between the warmest and coolest summers during the cycle was determined (Fig. 3). Sensitivity to the precession cycle is greatest over the Eurasian and North American continents because of their low heat capacity and large area. In the tropics the largest thermal response is seen over Africa. For an eccentricity of 0.04 (0.06), patterns of the thermal response are the same as in Fig. 2 with the isotherms re-scaled by a factor of 2 (3). Within the range of Earth's orbital eccentricity variations, the tropical thermal response to the precession cycle can be as large as 6°C over Africa. In the range of tropical temperatures this translates into a 50% change in

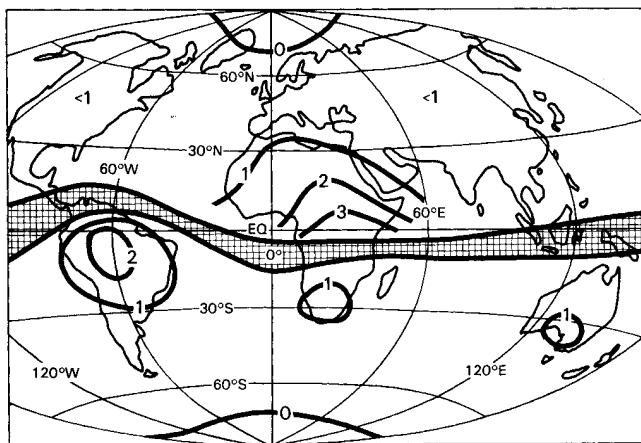


Fig. 4 Time (in kyr after summer solstice-perihelion alignment) of warmest summers during a precession cycle. The hatched area separates regions dominated by Northern and Southern Hemisphere summers.

saturation vapour pressure, implying a significant orbital impact on the tropical hydrological cycle.

The time of warmest summer temperatures during the 22-kyr precession cycle was determined and expressed as thousands of years past solstice-perihelion alignment (Fig. 4). Over the large tropical continent of Africa MST responds with lags of up to 3.5 kyr. Similar lags were seen in the maximum temperature contrast between North Africa and the surrounding tropical oceans. This demonstrates that an important thermal process related to the summer monsoon circulation is responding with a significant lag in the EBM. It may be that most of the several-thousand-year lag observed in tropical lake levels can be explained as resulting from a direct thermal response to the astronomical forcing.

Time lags of MST over Eurasia and North America are <1 kyr (Fig. 4). This is consistent with the fresh melt water signal that is observed in the North Atlantic⁹, relating to the wastage of continental ice sheets. However, it presents a contradiction between our EBM results and interpretation of the upwelling signal observed in the Arabian Sea². Prell's work² has employed the connection between radiative heating of the Asian continent (especially the Tibetan Plateau) during summer and development of the low-level Somali jet to interpret ocean core records in the Arabian Sea. A phase lag of $\frac{1}{4}$ cycle (5.5 kyr) between the upwelling signal and northern summer solstice-perihelion alignment is evident in the sedimentary record. Precession cycle experiments with GCMs may provide evidence of the relative importance of the Asian and African continents in determining the phase of the tropical climatic response in this region.

As expected, MST over southern continents peaks when the southern summer solstice is near perihelion. The small amplitude suggests that it may be difficult to detect a Southern Hemisphere phased response to the precession cycle in the climate record. Palaeoclimatic evidence suggests that Southern Hemisphere glaciers responded synchronously with their Northern Hemisphere counterparts during the past 25 kyr (ref. 10). This also seems to be the case for tropical precipitation/evaporation-controlled environments in the Southern Hemisphere⁶. Some components of the global environment (such as ice volume and sea level) are strongly modulated at the precession frequency, with a Northern Hemisphere phase because of the predominating continental response over middle and high latitudes. Our model results suggest that the primary tropical response to the precession cycle, variation of land-sea thermal contrast, is also driven with a Northern Hemisphere phase because of the large extent of tropical Africa in the Northern Hemisphere.

The encouragement and comments of T. J. Crowley and G. R. North are greatly appreciated.

Received 4 March; accepted 16 May 1986.

1. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121-1132 (1976).
2. Prell, W. L. in *Milankovitch and Climate Pt 1* (eds Berger, A. L. et al.) 349-366 (Reidel, Dordrecht, 1984).
3. Berger, A. L. *J. Atmos. Sci.* **35**, 2362-2367 (1978).
4. Kutzbach, J. E. & Guetter, P. J. in *Milankovitch and Climate Pt 2* (eds Berger, A. L. et al.) 801-820 (Reidel, Dordrecht, 1984).
5. Kutzbach, J. E. & Street-Perrott, F. A. *Nature* **317**, 130-134 (1985).
6. Street, F. A. & Grove, A. T. *Quat. Res.* **12**, 83-118 (1979).
7. North, G. R., Mengel, J. G. & Short, D. A. *J. geophys. Res.* **88**, 6576-6586 (1983).
8. Crowley, T. J., Short, D. A., Mengel, J. G. & North, G. R. *Science* **231**, 579-584 (1986).
9. Ruddiman, W. F. & McIntyre, A. *Science* **212**, 617-627 (1981).
10. Porter, S. C. *Quat. Res.* **16**, 263-292 (1981).

Crystallization microstructure in transparent monotectic alloys

William F. Kaukler* & Donald O. Frazier†

* Department of Chemistry, The University of Alabama in Huntsville, Huntsville, Alabama 35899, USA

† Space Science Laboratory, NASA Marshall Space Flight Center, Huntsville, Alabama 35812, USA

Transparent binary metallic alloy models are important in material and metallurgical science, as phase transformations and processes can be observed during solidification. The best transparent models for metallic growth and solidification morphology studies are members of the 'plastic' crystals group. These compounds have low entropies of fusion and cubic crystal structures, resulting in nonfaceted solid/liquid interfaces^{1,2}. We describe here the surprising morphology which occurs at the liquid₁/liquid₂/solid ($L_1/L_2/S$) triple junction in directionally solidifying miscibility gap systems at the monotectic temperature. The monotectic temperature in a binary mixture is the temperature at which two immiscible solutions of the same two components form phases in equilibrium with the solid phase of one of the components. These observations were made on succinonitrile (sn)-based systems ($C_4H_4N_2$), the 'plastic' crystal component, with a second component on a thermal gradient microscope apparatus³⁻⁵. Figure 1 shows two succinonitrile-based monotectic phase diagrams used in this study. The specimen cells contained sample gaps ranging from 50 to 150 μm , temperature gradients from 5 to 70 $^\circ\text{C cm}^{-1}$, and growth rates from 0.75 to 10.0 $\mu\text{m s}^{-1}$.

Part of the solidification front is composed of thousands of rapidly growing, cylindrical, curved hollow stalks of solid sn. The solidification process consists of the conversion of L_1 phase to sn + L_2 . The stalk interior contains L_2 phase, which protrudes above the stalk tip as shown in Figs 2 and 3. This droplet often detaches from the stalk, migrates up the thermal gradient, and dissolves. Together these features appear 'fungus-like' and grow in a variety of directions. The sn/water system sequence (Fig. 4a, b) demonstrates some of these dynamics. With the sn/water monotectic system, a second type of interface leads the other by 20-80 μm , and grows on the inner surface of the glass cell walls. This interface is of the planar type and grows in a coupled fashion, with the L_2 phase forming rods or channels in the solid sn matrix. Interphase spacings are typically 10 μm and are similar to those formed in metal monotectics grown in similar conditions. The fungus-like interface grows in the bulk solution, away from the glass cell walls. Cell gap thicknesses <20 μm suppress the fungus-like growth and enhance the planar interface.

Growth of an individual fungus-like structure gives the appearance that the L_2 head is leading, with the stalk following. Increasing the growth rate causes the dimensions of the features

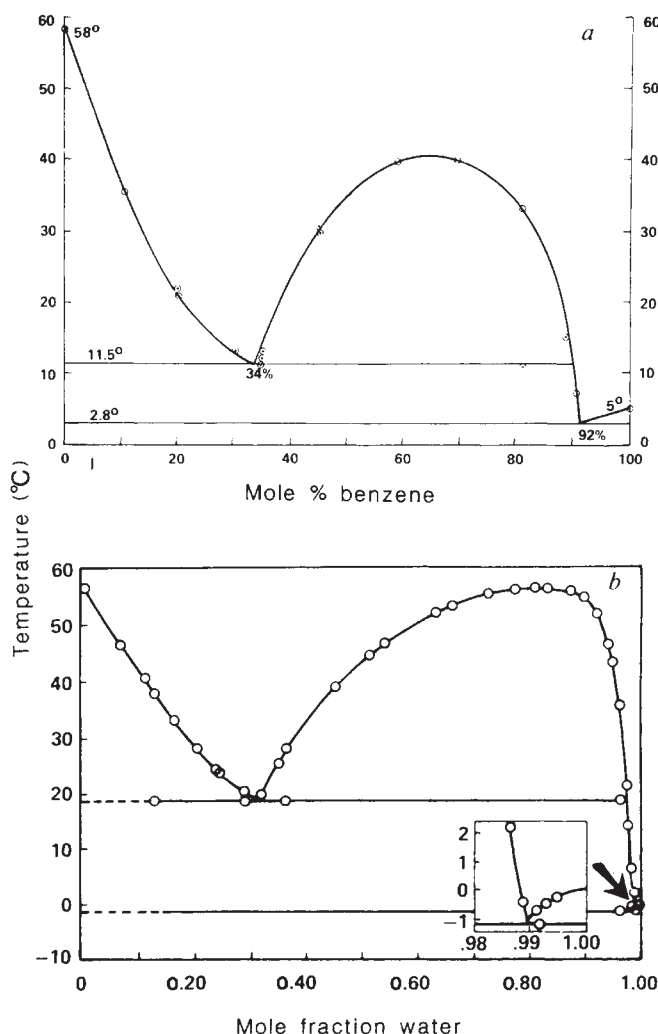


Fig. 1 Phase diagrams: *a*, succinonitrile/benzene; *b*, succinonitrile/water.

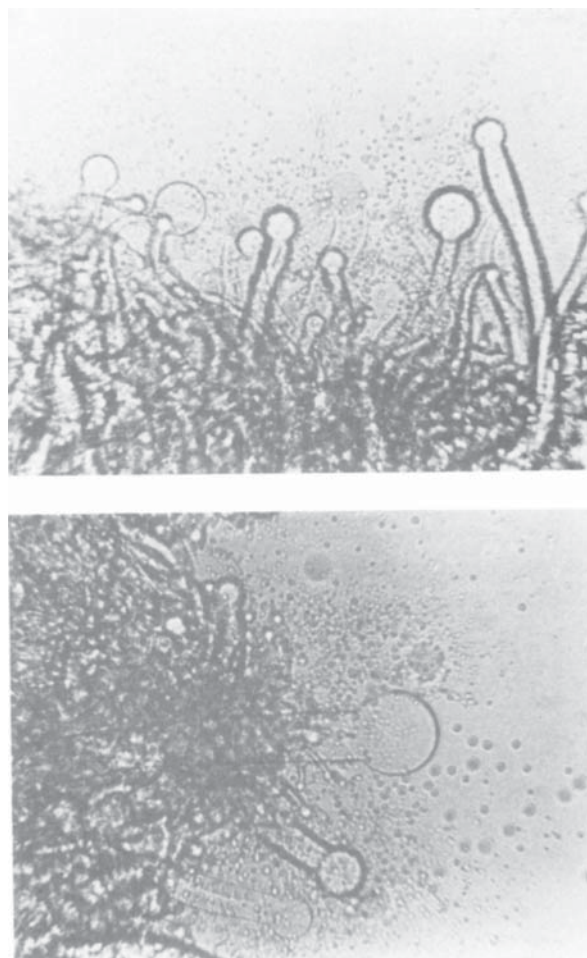


Fig. 2 Fungus-like structure of the succinonitrile/benzene solid/liquid interface. The larger features show detailed morphology better than in succinonitrile/water systems, including varying growth directions. Scale bars, 100 μm .

to diminish in proportion to the rate of increase of growth. This implies a kinetic limitation on the monotectic phase transformation at high growth rates, perhaps a direct dependence on the diffusion coefficient.

There is a similarity between this morphology and a known crystal growth morphology. Fine, single-crystal whiskers can be grown from a vapour by the vapour/liquid/solid (VLS)⁶ method. In this technique, a microscopic liquid droplet of metal insoluble in the solid, such as gold in the case of Si whiskers, enhances the condensation of Si vapour onto the substrate. Soon, a small crystal is deposited from the droplet onto the substrate, raising the droplet into the atmosphere. The droplet surface continues to enhance the condensation of the vapour and causes the single crystal whisker to grow and lengthen, all the time pushing the droplet along at its tip. This method increases the rate of crystal growth by ≥ 20 times over the rate for growth from the vapour. A similar mechanism, from a saturated liquid phase instead of a vapour phase, could apply in the monotectic case.

Although it is natural to review crystal growth mechanisms to find an analogous morphology, no known inorganic process is capable of fully explaining the formation of this particular monotectic morphology. Whereas the VLS crystal growth technique is viably analogous, arguments which explain droplet formation at the stalk tip on the basis of large internal pressure increases inside the stalk, at the instant of monotectic conversion, are even more viable and find similarities in the world of mycology. Excellent biological models may be found in the

growth of some filamentous fungi such as *Rhizopus nigricans* or *Phycomyces blakesleeana*. Similar mechanisms of shape acquisition, such as internal pressure effects, can thus exist in two otherwise unrelated substances such as plants and crystals.

Just as the stalks appear to be tubes of sn filled with the L_2 phase at their core, which protrude to form heads on the stalk, some moulds have contiguous spore head and stalk contents without septa. Of the ways by which sporangia may separate from sporangiophores, some may easily apply to monotectics. One such discharge mechanism in fungi employs turgor pressure within the basidium to round off the cell walls, causing spore separation. In the analogous case for monotectic solidification, high internal stalk pressure may result from the density change of a volume element converted from L_1 to equilibrium L_2 and pure sn. Dilatometric measurements show that at the instant of conversion, nearly 'pure' L_2 phase is separated from nearly 'pure' sn. If this occurs in the concentration region where sn is the majority phase, an interfacial free energy minimum will dictate that the newly formed 'pure' sn phase will insert itself between the sn-rich matrix and the newly formed L_2 phase. Formation of individual tubules spaced by the matrix phase is probably a consequence of exothermic heat transfer, like that occurring in dendritic growth. In this morphological configuration, the former L_1 phase has converted to an sn solid stalk of relatively high density (low volume) surrounding a low-density (high-volume) fluid core of L_2 phase. The result is that the internal pressure must be initially quite high, causing extrusion

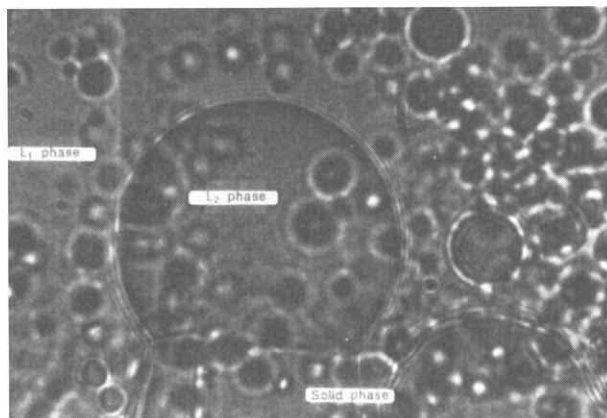


Fig. 3 Succinonitrile/benzene 'fungus' morphology at high magnification ($\times 960$). L_1 phase, succinonitrile-rich; L_2 phase, water-rich; solid phase, 'pure' succinonitrile.

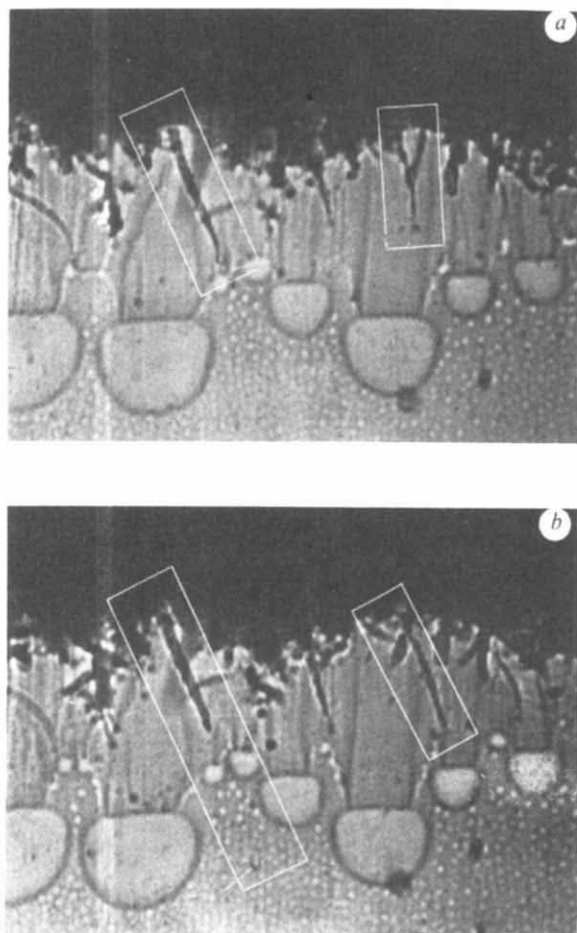


Fig. 4 Structures at a succinonitrile/water interface: **a**, Initial positions of head (L_2 phase) and tubular structure ('pure' succinonitrile stalk). **b**, L_2 head of fungus-like structure has popped off and traveled up the thermal gradient in the L_1 matrix.

of some L_2 phase from the core, followed by L_2 phase expulsion, possibly enhanced by the thermal gradient. Furthermore, separation of the head from the stalk is consistent with minimization of interfacial areas associated with the triple junction by rounding of the stalk and head. Of course, confirmation will depend on further experiments for determining composition profiles.

Existing knowledge forces us to accept that the solidification process involves thermodynamic and kinetic effects, which act

together to yield a variety of phase distribution patterns⁷. Much depends on the relative densities of phases, interfacial tension variations with temperature and composition, thermal conductivities, certain surface affinities, activity coefficients, general fluid dynamics, and so on. There are many parameters which act together to yield the metallographic pattern observed in a cross-sectioned metal ingot.

We acknowledge the support of NASA and the Universities Space Research Association. We appreciate helpful discussions with Professors Harold Wilson and Jeanette Jones, and the support of Dr A. J. Dessler.

Received 30 November 1985; accepted 2 April 1986.

1. Timmermans, J. *J. phys. chem. Solids* **18**(1), 1-8 (1961).
2. Jackson, K. A. & Hunt, J. D. *Acta Metall.* **13**, 1212-1215 (1965).
3. Kaukler, W. F. *Rev. Sci. Instrum.* **55**, 1643-1647 (1984).
4. Kaukler, W. F. thesis, Univ. Toronto (1981).
5. Hunt, J. D., Jackson, K. A. & Brown, H. *Rev. scient. Instrum.* **37**, 805 (1966).
6. Wagner, R. S. & Ellis, W. C. *Trans. metall. Soc. A.I.M.E.* **233**, 1053-1064 (1965).
7. Frazier, D. O., Facemire, B. R. & Fanning, U. S. *Acta Metall.* **34**, 63-72 (1986).

The curvature of Wadati-Benioff zones and the torsional rigidity of subducting plates

Michael Bevis

Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina 27695-8208, USA

It is an axiom of plate tectonics that lithospheric plates are to a good approximation torsionally rigid, that is, inextensible in the 'plane' of the plate¹⁻³. Although few scientists would question this assumption when it is applied to oceanic lithosphere at the Earth's surface, there is no consensus as to whether or not oceanic lithosphere maintains its mechanical integrity during subduction (compare refs 4-7 with refs 8-11). I show here that the shapes of many Wadati-Benioff zones are inconsistent with the notion that they manifest subducted portions of oceanic plates subject only to very small ($\leq 1\%$) membrane strains.

If the subducted and surface portions of a plate do contain a coherent and continuous elastic layer of nearly constant torsional rigidity then a subducting plate should deform in approximately the same manner as a perfectly flexible but perfectly inextensible surface. This is because the displacements of a surface which is only slightly extensible differ from those of an inextensible surface by quantities which are of the same order of magnitude as the extensibility of the surface¹². Thus if we assume, for example, that the membrane strains in a subducting plate are $< 1\%$, and we consider the shape of a slab with a lateral extent of 1,000 km, then a surface centred along the plate will deform such that its displacement will differ from that of a flexible but inextensible surface by ≤ 10 km. This discrepancy is less than the thickness of Wadati-Benioff Zones (WBZs) and the subducted plates that they are presumed to manifest. We must then consider whether the observed shapes of WBZs (or of trend surfaces fitted⁷ to these hypocentre clouds) are consistent with the suggestion that they can be modelled by a flexible but inextensible surface that can characterize also an oceanic plate at the Earth's surface.

At any point on a surface there are among the normal sections one for which the surface curvature is a maximum and one for which it is a minimum; the directions of these sections are at right angles to each other. These extreme curvatures, C_1 and C_2 , are known as the principal curvatures, and the corresponding radii $R_1 = 1/C_1$ and $R_2 = 1/C_2$ are the principal radii of curvature. Convex curvatures are positive, and concave curvatures are negative. The gaussian curvature, K , at any point on the

surface is given by $K = C_1 C_2 = 1/R_1 R_2$. A well-known theorem of differential geometry due to Gauss states that a flexible but inextensible surface transforms such that the gaussian curvature at every point remains unchanged¹³.

The gaussian curvature of a spherical cap of radius r is $K_c = 1/r^2$ at every point on the surface. We can characterize an oceanic plate at the Earth's surface by choosing r such that the cap will bisect (longitudinally) the elastic lithosphere. We take $r = 6,354$ km, and find $K_c = 2.477 \times 10^{-8} \text{ km}^{-2}$. We must consider whether it is normally possible to characterize the shape of a WBZ with a surface whose gaussian curvature is everywhere K_c . (Note that the 'indented ping-pong ball' model introduced by Frank⁶ invoked an axially symmetric subduction geometry in which the slab is characterized by a concavo-concave surface with $R_1 = R_2 = -r$ and $K = 1/r^2 = K_c$. This is a specific example of a model geometry consistent with the general requirement that the model surface is everywhere characterized by $K = K_c$.)

An examination of extant contour maps and sections depicting the regional configuration of WBZs indicates that in many areas there are major departures from the shapes attainable by a flexible and inextensible spherical cap. For example, in the depth range 50–150 km the Marianas WBZ has its maximum curvature along a near-vertical section¹⁴ (roughly down-dip) with $R_1 \approx 200$ km. To preserve a gaussian curvature $K_c \approx 2.5 \times 10^{-8} \text{ km}^{-2}$ we would require $R_2 \approx 2 \times 10^5$ km, that is, only a very slight curvature along a horizontal section. In fact, the Marianas WBZ is strongly curved along strike¹⁴, and $R_2 \approx -550$ km.

It is not, in fact, necessary to engage in such arithmetical exercises to arrive at the main conclusion. An inextensible spherical cap deforms so as to maintain its positive gaussian

curvature at every point, so it can never assume a concavo-convex shape (which implies $K < 0$). Subducting plates typically attain dips of a few degrees at the trench axis, 10° – 20° in the interplate thrust zone, and 30° – 65° at depths near 75 km; that is, throughout this depth range essentially all subducting plates are convex along a vertical section¹⁴. Thus the map-view shapes¹⁵ of the trench axes, the shallow interplate seismic zones and the shallowest portions of the WBZs associated with the Alaska–Aleutian, Sumatra–Java–Flores, Caribbean, Scotia, Ryukyu and Marianas arcs indicate that the uppermost portions of the oceanic plates subducting in these areas are concavo-convex in shape.

On purely geometrical grounds it seems that the assumption of near torsional rigidity commonly applied to surface oceanic plates is not generally applicable to their subducted extensions.

Received 4 April; accepted 27 May 1986.

1. McKenzie, D. P. & Parker, R. L. *Nature* **216**, 1276–1280 (1967).
2. Morgan, W. J. *J. geophys. Res.* **73**, 1959–1982 (1968).
3. Le Pichon, X. *et al. Plate Tectonics* (Elsevier, Amsterdam, 1973).
4. Elsasser, W. M. in *The Application of Modern Physics to Earth and Planetary Interiors* (ed. Runcorn, S. K.) 223–246 (Wiley, New York, 1967).
5. Isacks, B. L. *et al. J. geophys. Res.* **73**, 5855–5899 (1968).
6. Frank, F. *Nature* **220**, 363 (1968).
7. Bevis, M. & Isacks, B. L. *J. geophys. Res.* **89**, 6153–6170 (1984).
8. Ringwood, A. E. & Green, D. H. *Tectonophysics* **3**, 383–427 (1966).
9. Kanamori, H. *Tectonophysics* **12**, 187–198 (1971).
10. Stöiber, R. E. & Carr, M. J. *Bull. volc.* **37**, 304–325 (1973).
11. Freund, R. *et al. Mechanisms of Continental Drift and Plate Tectonics* (eds Davies, P. A. & Runcorn, S. K.) 17–39 (Academic, London, 1980).
12. Jellett, J. H. *Trans. R. Ir. Acad.* **22**, 343–377 (1855).
13. Stoker, J. J. *Differential Geometry* (Wiley, New York, 1969).
14. Isacks, B. L. & Barazangi, M. in *Island Arcs, Deep Sea Trenches and Back-Arc Basins* (eds Taiwani, M. & Pitman, W.) 99–114 (American Geophysical Union, Washington, 1977).
15. Espinosa, A. F. *et al. Seismicity of the Earth 1960–1980* (Marie Tharp, Oceanographic Cartographer, South Nyack, 1981).

A deep seismic reflection profile across the northern North Sea

Alastair Beach

Midland Valley Exploration Limited, 14 Park Circus,
Glasgow G3 6AX, UK

The Viking Graben is a major Mesozoic rift basin in the northern North Sea¹. Seismic refraction data², gravity surveys³ and theoretical modelling⁴ have led to the general acceptance of an extensional tectonic origin for the Viking Graben, involving an early Mesozoic phase of extension and fault-controlled subsidence, followed by a more regional, thermal-contraction subsidence. The early success of the BIRPS deep seismic reflection profiling⁵ and the results from COCORP⁶ led to the planning and acquisition during 1983 of a deep seismic line, recorded to 15 s TWT (two-way time), across the north Viking Graben. The aims of the present study were to provide constraints on the geometry of the Viking Graben and its extensional faults, to provide data on the geometry of crustal thinning beneath the graben, and to provide an insight into the relation between the shallow extensional features of the sedimentary basin and the deep crustal thinning. Here we present a specific interpretation of this seismic profile; further details will be given elsewhere⁷.

The seismic profile was shot from north of the Shetland Islands to the coast of Norway near Bergen, a length of ~340 km, in a NW–SE direction. The line crossed what was considered to be the structurally simplest part of the Viking Graben; its location is shown in Fig. 1. There is a striking similarity between the extensional geometry of the Viking Graben and that embodied in a model by Wernicke⁸, based on the Basin and Range province. Such an interpretation has implications for intraplate crustal deformation related to tectonic activity at plate boundaries, because low-angle detachments within the lithosphere will link

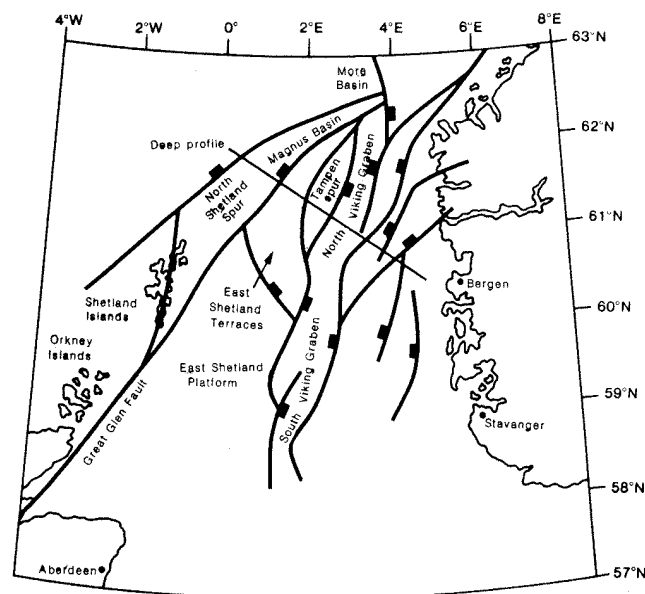


Fig. 1 Location map of the deep profile across the northern North Sea.

crustal tectonic displacements over very large and widely separated areas. There is a marked difference between the geometry seen on the deep profile and current structural models for the Viking Graben based on conventional seismic data¹.

The line drawing of the profile is shown in Fig. 2a. The interpretation of the profile can be considered in three stages. First, the upper part of the section, occupied by Mesozoic and Tertiary sequences, is well known from the results of hydrocarbon exploration¹. The clarity of the reflections allows major fault blocks to be defined (Fig. 2b); the principal unconformity shown separates tilted Jurassic (and possibly older) sequences

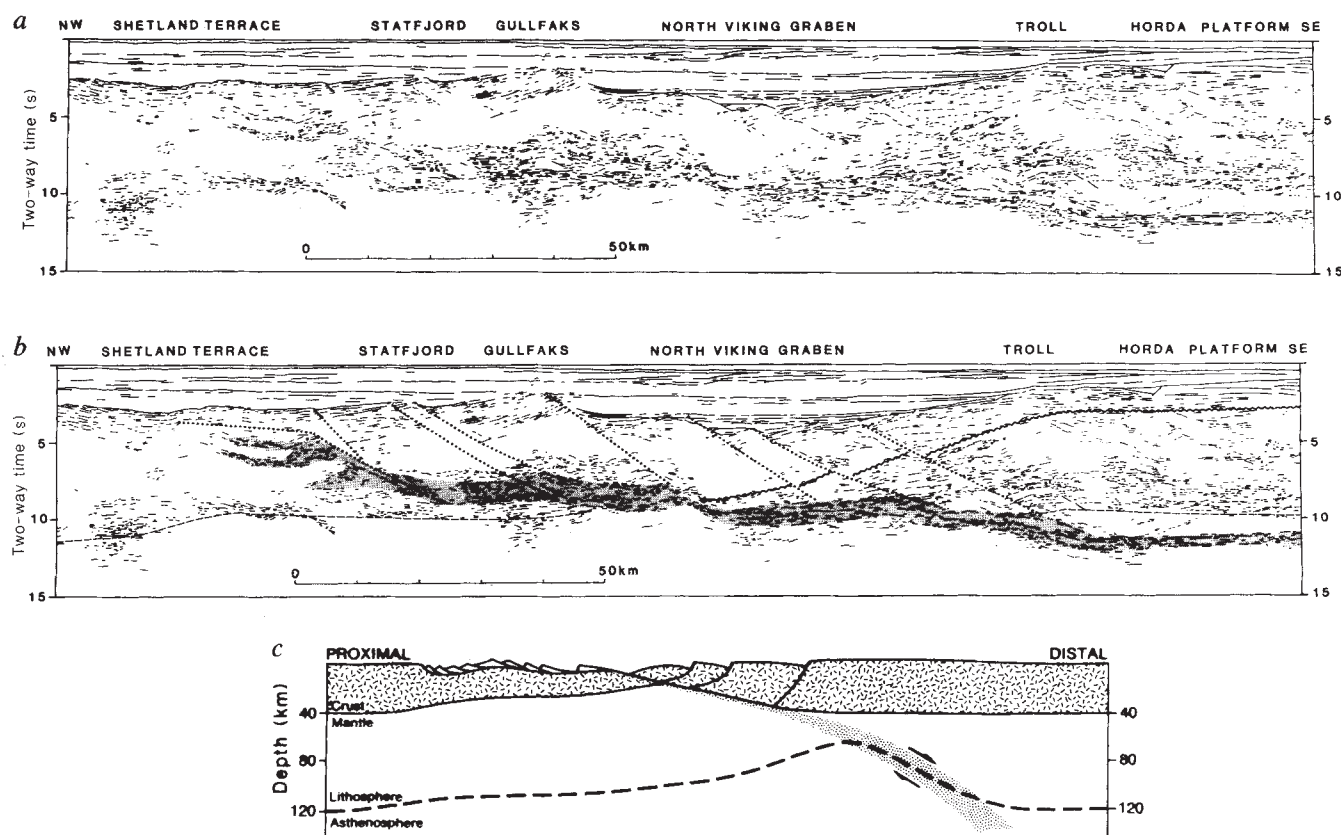


Fig. 2a, Line drawing of profile, picking out reflector strength and continuity. Horizontal = vertical scale where seismic velocity = 5 km s^{-1} . Above this velocity the vertical scale is greater; below it, the horizontal scale is greater. Redrawn from ref. 7. **b**, Interpretation of **a**. The shaded area represents a major extensional shear zone passing upwards from the mantle in the east to the middle crust in the west. Major eastward-dipping extension faults, affecting the hanging wall of the shear zone, are shown as dotted lines. The base of the crust is shown as a dashed line; the wavy line, drawn as a boundary to areas of different seismic character, reveals a large-scale crustal rollover in the hanging wall. In the eastern portion of the profile the wavy line corresponds approximately to the base of the Devonian, and top of the metamorphic basement. **c**, Geometrical model for crustal extension, redrawn from ref. 8.

from flat-lying Cretaceous sequences. The upstanding and tilted block Gullfaks is finally buried by end-Cretaceous times (Fig. 2b). Second, the base of the crust is interpreted to lie below the zone of strong seismic reflections, by analogy with other areas around the United Kingdom where combined seismic reflection and refraction data have identified this as corresponding to the Moho⁹. Third, the area between the lower crust and the base of the sedimentary sequences is rather featureless in terms of seismic reflections. This part of the section may be occupied by the sediments of a Devonian basin (the edges of which are seen in north-east Scotland and on the Norwegian coast around Bergen), overlying Caledonian metamorphic rocks.

The most notable geometrical aspect of the profile is the asymmetric disposition of the seismic reflections: six different features testify to the asymmetry of the geometrical development of the graben (Fig. 2b). First, the Viking Graben is an asymmetric half-graben filled with a westerly-thickening wedge of Cretaceous sediments. Second, the fault blocks of tilted Jurassic sequences show a consistent westerly tilt. Third, the disposition of these blocks is asymmetric—they are present in the west from the Shetland Terrace to the floor of the Viking Graben, but absent to the east. Fourth, the principal interpreted extensional faults in the upper crust all dip to the east. Fifth, the seismic reflections in the lower half of the crust are weakest in the west and become progressively stronger, as well as deeper, to the east. Sixth, the most continuous and strongest seismic reflections are present at ~12 s TWT below the Horda Platform, well to the east of where the maximum upper-crustal extension is seen.

The essentials of Wernicke's geometrical model⁸ for crustal extension are shown in Fig. 2c. There is a marked similarity

between the overall geometry of the Viking Graben profile and the Wernicke model, and the sense of asymmetry described above is interpreted to result from a low-angle extensional detachment dipping east under the Viking Graben. The Norwegian part of the profile shows a huge crustal rollover at the base of the Mesozoic, from which an eastward-flattening sole detachment is deduced; that is, the hanging-wall geometry results from accommodation to sole-fault shape.

The lower-crustal zone of reflections is interpreted to result from a major crustal extensional shear zone. It passes from the lower crust to the base of the crust beneath the western margin of the Viking Graben, and from the base of the crust to the upper mantle beneath the western edge of the Horda Platform. An upper-crustal hinge is developed above the location where the detachment is interpreted to pass into the mantle, shown in Fig. 2b as modelled by Wernicke. The deep reflections at 12 s below the Horda Platform are interpreted to arise from the major detachment shear zone within the upper mantle. Interpretations of seismic refraction and gravity data^{2,3} indicate that the Moho in this region lies at a depth of ~30 km, or ~10 s TWT. A more detailed seismic refraction profile on the Horda Platform (Elf Aquitaine, unpublished data) supports the interpretation of the Moho at 10 s TWT. Thus, on the profile presented here, the Moho has been drawn at the base of the reflective zone at 10 s (see Fig. 2b) and the deeper, strong reflections are therefore within the upper mantle.

Wernicke's model for crustal extension explains the principal geometrical features of the Viking Graben profile, and is thus a reasonable model for the development of the Viking Graben. Using the simple velocity structure obtained from seismic refraction

tion², the position of the extensional detachment in approximate depth conversion can be estimated. It lies at nearly 40 km below the Horda Platform, and deforms the base of the crust at 20–25 km depth, rising westwards through the crust to ~15 km depth. The interpretation of a low-angle extensional detachment beneath a rift provides a mechanism for the linking of the tectonic displacement of the crust over very large distances, and enables a relationship to be established between the deformation of plate interiors and activity at plate boundaries. It is more likely that the low-angle extensional detachment soles out at the base of the lithosphere, into the weak, low-velocity layer of the upper asthenosphere. There may be decoupling of deformation between the lithosphere and the asthenosphere. The presence of the weak, low-velocity layer permits rapid transmission of displacements from the plate edge to the plate interior in response to plate boundary forces. In the case of the Viking Graben, the southeastward-dipping detachment is interpreted to be bounded by strike-slip displacements on the Tornquist Zone, which runs south-east from Denmark, through Poland, to the Tethyan margin east of the present Carpathians^{1,10}. Such a linkage would transfer displacements over a distance of ~1,800 km from the plate boundary to the zone of intraplate extension. With the lithosphere deforming and displacing as a hanging wall above an upper-asthenosphere detachment, major tectonic structures in the crust such as the Viking Graben and the Tornquist Zone define the scale over which deformation occurred. The deep seismic profile across the Viking Graben fills part of the major interpretative gap between studies of plate motions and field studies of deformation and displacement of upper-crustal rocks.

Following the Wernicke model⁸, regional subsidence patterns are related to broad changes in the isostatic balance of crustal and mantle material. Maximum syntectonic subsidence occurs above the zone of maximum crustal thinning, as more dense mantle material occupies the position of crustal material, and a graben is created. Once the detachment is contained within the upper mantle, only mantle lithosphere is thinned. As mantle lithosphere is more dense than the asthenosphere, thinning the former and replacing it by the latter causes uplift during extension. Thus, east and south-east of the Viking Graben–Horda Platform hinge line it is suggested that syntectonic uplift occurred. Thermal subsidence is superimposed on the tectonic subsidence of the Viking Graben as the upwelled mantle lithosphere cools. Thermal subsidence also occurs over the area of tectonic uplift, as the upwelled asthenosphere cools, and, broadly speaking, there is no net resultant subsidence in this region. Thus, the subsidence profile across the Viking Graben is generally consistent with the Wernicke model of a southeastward-dipping detachment. Crustal extension occurred across the Shetland Terrace and as far east as the position of Troll on Fig. 2b. Mantle-lithospheric extension occurred from below Gullfaks southeastwards, below the Viking Graben and the Horda Platform. Mantle-only extension occurred from beneath the position of Troll southeastwards, below the Horda Platform and then below Norway. The upper crustal hinge at the position of Troll on Fig. 2b is thus an important regional tectonic feature.

The geometry of the Viking Graben and its development as the hanging wall of a major low-angle extensional detachment has great economic importance, as the sizes of structure and the sizes of hydrocarbon fields vary systematically from west to east. The Shetland Terrace may be typified by Thistle Field, with an estimated 10⁹ barrels of oil initially in place¹¹. To the east, the two large structures of Statfjord and Gullfaks are characterized by an estimate for Statfjord of 5.6 × 10⁹ barrels of oil initially in place¹². The east flank of the Viking Graben is characterized by much larger hydrocarbon accumulations; for example, Troll Field, with an estimated 4.6 × 10¹³ standard cubic feet of gas in place¹³ (7.5 × 10¹⁰ barrels oil equivalent), is situated at the upper-crustal hinge point discussed above. The asymmetry in hydrocarbon field sizes emphasizes the asymmetric tectonic development

of the Viking Graben. The smaller extensional fault-block structures developed at the north-west or proximal end, where shallow to mid-crustal detachment of extensional faults occurred. Larger fault blocks developed (Statfjord, Gullfaks) where deep to base-crustal detachment of extensional faults occurred. The largest hydrocarbon accumulations relate to the much larger scale of development of the hanging-wall rollover and crustal hinge on the south-east or distal side of the Viking Graben, that is, the Troll area. Thus, the development and refinement of models of crustal extension using deep seismic data is important for understanding the geographical distribution of hydrocarbon reserves.

Received 12 March; accepted 3 June 1986.

1. Ziegler, P. A. *Geological Atlas of Western and Central Europe* (Shell International London, 1982).
2. Solli, M. thesis, Univ. Bergen (1976).
3. Donato, J. A. & Tully, M. C. in *Petroleum Geology of the Continental Shelf of NW Europe* 65–75 (Heyden, London, 1981).
4. McKenzie, D. P. *Earth planet. Sci. Lett.* **40**, 25–32 (1978).
5. Smythe, D. K. *et al. Nature* **299**, 338–340 (1982).
6. Smithson, S. B., Brewer, J. A., Kaufman, S., Oliver, J. E., & Hurlrich, C. E. *J. geophys. Res.* **84**, 5955–5972 (1979).
7. Beach, A., Bird, T. & Gibbs, A. D. in *Extensional Tectonics* (Geological Society, in the press).
8. Wernicke, B. *Can. J. Earth Sci.* **22**, 108–125 (1985).
9. Barton, P., Matthews, D., Hall, J. & Warner, M. *Nature* **308**, 55–56 (1984).
10. Pegrum, R. M. in *Petroleum Geology of North European Margin*, 359–370 (Graham & Trotman, London, 1984).
11. Hallet, D. in *Petroleum Geology of the Continental Shelf of NW Europe* 315–325 (Heyden, London, 1981).
12. Kirk, R. H. in *Giant Oil and Gas Fields of the Decade 1968–1978*, 95–110 (American Association of Petroleum Geologists, Tulsa, 1980).
13. *Oil Gas J.* **83**, 41 (1985).

Helium flux in a continental land area estimated from ³He/⁴He ratio in northern Taiwan

Yuji Sano*, Hiroshi Wakita* & Chin-Wang Huang†

* Laboratory for Earthquake Chemistry, Faculty of Science, University of Tokyo, Bunkyo-ku, Tokyo, Japan

† Department of Chemistry, Chung-Yuan Christian University, Chung-Li, Taiwan, Republic of China

The helium flux and its isotopic composition may provide useful information not only on the helium budget in the atmosphere but also on the uranium and thorium contents of the rocks in the Earth's crust and mantle and the generation of heat from radioactive decay of these elements. Data have been available from oceanic areas^{1–4}, but those from continental areas are sparse^{5–7}. We report here the ³He and ⁴He fluxes derived from measurement of the ³He/⁴He ratio of natural gases in the northern part of Taiwan. The estimated ³He and ⁴He fluxes are 3.9 and 2.7 × 10⁶ atoms cm⁻² s⁻¹ at the Chinshui site, and 7.2 and 2.4 × 10⁶ atoms cm⁻² s⁻¹ at the Chuhuankeng site. The ³He fluxes are in approximate agreement with the global flux estimated from excess ³He in sea water¹. The ⁴He fluxes are comparable to the ⁴He flux⁶ for the continents deduced from terrestrial heat flow data and the heat/He flux ratio, but are significantly larger than those reported for the ocean floor. The relationship observed between the He isotope ratio and CO₂ and CH₄ concentrations suggests that there is a mixing trend between the high ³He/⁴He–high CO₂–low CH₄ component and the low ³He/⁴He–low CO₂–high CH₄ component. The former may be derived from the upper mantle and the latter produced in a sedimentary environment.

In spite of the ease with which helium escapes from the upper part of the atmosphere to interplanetary space because of its low atomic mass, the terrestrial atmosphere contains an appreciable amount of helium with a concentration of 5.24 × 10⁻⁶ by

Table 1 Chemical compositions and $^3\text{He}/^4\text{He}$ and $^4\text{He}/^{20}\text{Ne}$ ratios in natural gases from Taiwan

Site	Depth* (m)	CH ₄ (%)	C ₂ H ₆ (%)	CO ₂ (%)	N ₂ (%)	He (p.p.m.)	Ar (p.p.m.)	$^3\text{He}/^4\text{He}$ ($\times 10^{-6}$)	$^4\text{He}/^{20}\text{Ne}$
Chinshui									
a	2,750	92	4.9	2.9	0.41	24	34	1.63 ± 0.04	22
b	3,390	91	3.2	5.7	0.43	43	<20	1.80 ± 0.06	2,700
c	3,460	86	3.2	9.7	0.34	55	<20	1.82 ± 0.06	2,200
d	3,575	84	3.4	8.3	0.28	65	25	1.74 ± 0.04	2,800
e	3,935	—	—	—	—	—	—	1.77 ± 0.05	14
f	4,045	90	2.8	6.7	0.87	48	720	1.79 ± 0.03	100
g	4,100	86	2.8	11	0.30	40	<20	1.84 ± 0.05	45
h	4,450	85	2.4	12	0.48	43	<20	1.89 ± 0.05	53
Chuhuangkeng									
a	2,035	69	1.5	28	1.3	80	49	3.29 ± 0.09	3,000
b	2,475	67	1.0	30	2.0	92	50	3.47 ± 0.12	1,100
c	2,885	55	0.85	43	1.5	90	<20	3.77 ± 0.13	3,300
d	3,260	55	0.96	42	1.7	88	<20	3.75 ± 0.05	3,200
e	3,600	—	—	—	—	—	—	3.68 ± 0.08	1,000
Air		—	—	0.04	78.1	5.2	9,340	1.40	0.318

* Depth of the gas reservoir.

volume. This has been attributed to a supply from the outgassing of the Earth's crust and mantle. There may be an approximate balance between input from the crust and mantle and output at the high-temperature exosphere. Based on the He budget in the atmosphere and the generation of heat from radioactive decay, the He flux from the solid Earth has been estimated by many investigators, although mainly for oceanic areas. Such data as are available for land areas were obtained indirectly, by assuming parameters such as total inventory of CO₂, He/CO₂ ratio in a volcanic gas⁵, terrestrial heat flow data, and heat/He flux ratio⁶. Our new data on both ^3He and ^4He fluxes in a continental land are independent of these *a priori* assumptions.

The natural gas samples were collected at two commercial gas wells at varying depths in the western coastal plain of northern Taiwan. The sample container was made of lead glass with vacuum valves at both ends. At the sampling site, the container was filled with distilled water and one end was connected to the output of the natural gases. After the water in the container had been completely replaced by the gases, both valves were closed. Thus atmospheric contamination was kept to a minimum during sample collection. After the containers were brought back to the laboratory, the chemical compositions of the natural gases (CH₄, C₂H₆, CO₂, N₂, He and Ar) were analysed by gas chromatography⁸. Overall analytical margins of error were estimated to be <5% for most samples. Argon contents have larger margins of error due to their low abundance. $^3\text{He}/^4\text{He}$ and $^4\text{He}/^{20}\text{Ne}$ ratios were measured using a mass spectrometer. Experimental details, including purification of He and Ne and correction for a hot blank, mass discrimination and the resolving power of the mass spectrometer, have been given elsewhere^{9,10}.

Table 1 lists the results of chemical analyses and measurements of the $^3\text{He}/^4\text{He}$ and $^4\text{He}/^{20}\text{Ne}$ ratios in our samples, together with the depths of the gas reservoirs. The major constituent of most of the gas samples is CH₄. Significant amounts of CO₂ are observed in the Chuhuangkeng samples, although this site is situated on a thick sedimentary layer. There is a positive correlation between sample depth and CO₂ content and a negative correlation between depth and C₂H₆ content in the Chinshui samples. A similar tendency is found in the Chuhuangkeng samples. On the other hand, there is no correlation between depth and N₂, He and Ar contents.

It has been generally accepted that CH₄-rich natural gas is produced by the anaerobic bacterial decomposition of organic matter in an anoxic environment¹¹. Low $^3\text{He}/^4\text{He}$ ratios in natural gases from continental areas have been reported by

many investigators^{12,13}. These results were thought to support the biogenic origin of CH₄, as opposed to a magmatic origin. Recently, however, unexpectedly high $^3\text{He}/^4\text{He}$ ratios, up to 8.65×10^{-6} , were found in CH₄-rich natural gases in the 'Green Tuff' formation (volcaniclastic rocks formed in the middle Miocene) of Japan¹⁴. Given the presence of inorganic CH₄ accompanied by mantle-derived He at mid-ocean ridges, most Japanese natural gases have been attributed to a magmatic origin.

The observed $^3\text{He}/^4\text{He}$ ratios vary from 1.63×10^{-6} to 3.77×10^{-6} . All samples show ratios higher than the atmospheric value of 1.40×10^{-6} . The high $^3\text{He}/^4\text{He}$ ratio strongly suggests the contribution of mantle-derived He, which may result in the presence of magmatic volatiles in the samples. The relationship between the $^3\text{He}/^4\text{He}$ ratio and CH₄ concentration is shown in Fig. 1. A negative correlation is obvious for these data from two wells. Statistical significance is calculated separately: the correlation coefficient is -0.55 for the Chinshui well and -0.97 for the Chuhuangkeng well. This suggests that there is a mixing trend between two distinct components. One is a mixture of He with a high $^3\text{He}/^4\text{He}$ ratio and CO₂. The other is a mixture of He with a low $^3\text{He}/^4\text{He}$ ratio, CH₄ and C₂H₆. The first component may be derived from the upper mantle. Quaternary volcanism, however, is not found in Taiwan. The magmatic He and probably a part of the CO₂ may be derived either from the andesitic volcanism that occurred in the Pleistocene or from current plutonic activity related to the collision-type tectonics. The second component may be produced in the sedimentary environment, as suggested by its lower He isotope ratio. Hydrocarbons in these wells may be attributed to a biogenic origin.

The $^3\text{He}/^4\text{He}$ ratios of eight gas samples were measured at varying depths in the Chinshui well. The ratios increase with sampling depth. A similar tendency is observed at the Chuhuangkeng well. Figure 2 indicates the relation between the $^3\text{He}/^4\text{He}$ ratio and the depth. The tendency is interpreted as meaning that the He in these samples is a mixture of two sources: *in situ* radiogenic He in sediments, and mantle He leaking through gas reservoirs and sediments. The deeper the gas reservoir, the thicker is the sediment layer at the ground surface, which yields a larger radiogenic He contribution from the sediment. If a steady state of He flow is assumed in the vertical sediment column, which means He radiogenically produced in the sediment column is immediately released from production sites into the gas phase, the following equation can apply:

$$(^3\text{He}/^4\text{He})_t = \frac{FR_g - Pr_s}{F - Pr}$$

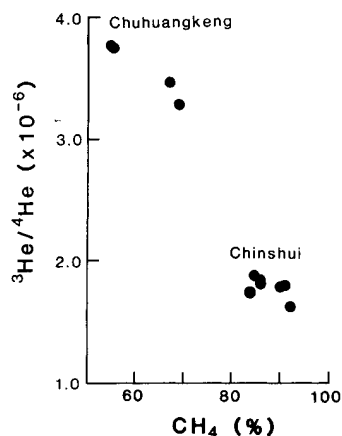


Fig. 1 Relation between $^3\text{He}/^4\text{He}$ ratios and CH_4 concentrations of natural gases at Chinshui and Chuhuangkeng sites.

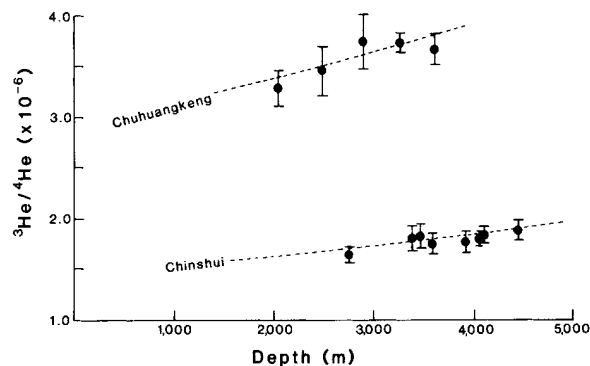


Fig. 2 Relation between $^3\text{He}/^4\text{He}$ ratios and sampling depths of natural gases at the Chinshui and Chuhuangkeng sites. Bars represent measurement error (± 2 s.d.). Dashed lines are best-fit curves of the equation derived from the steady-state flow of He in the vertical sediment column.

where t is the sample depth, F the ^4He flux at the ground surface at the site, P the ^4He production rate in the sediment, R_s the $^3\text{He}/^4\text{He}$ ratio at the surface, and R_s the $^3\text{He}/^4\text{He}$ ratio of radiogenic He in the sediment. As the thickness of the sediment in the area studied was estimated to be ~ 9 km by a seismological study¹⁵, t should be < 9 km.

There are no available data for the uranium and thorium contents of the sedimentary rock in the investigated area. Assuming typical U and Th contents of 3 and 10 p.p.m., and a rock density of 2.7 g cm^{-3} (ref. 16), the production rate of radiogenic ^4He is calculated to be $1.5 \text{ atoms cm}^{-3} \text{ s}^{-1}$. The *in situ* radiogenic production rate of ^3He may also depend on the chemical composition of the rock, especially on the Li content¹⁷. According to calculation¹⁸, the radiogenic $^3\text{He}/^4\text{He}$ ratio of typical crustal rocks is $\sim 2 \times 10^{-8}$, which agrees well with the ratio measured in fluids collected in ancient platforms. Taking the values of R_s (2×10^{-8}) and t (the depth of the gas reservoir), we can calculate the ^3He and ^4He fluxes by fitting observed values to the above equation. In the calculation a nonlinear least-squares method (Gauss-Newton method) is adopted. The best-fit curves (dotted lines in Fig. 2) are derived by the iterative improvement method.

The advantage of this method lies in its independence of the absolute concentration of He in the sample. In the measurement of He flux at the ocean floor, the vertical gradient of He concentration and the diffusion coefficient, or advection flow model, are generally taken into account. In contrast, the present method is not affected by the He concentration, which changes with variation in the production rate of CH_4 in the sediment; it depends only on the isotopic composition, which cannot be distorted by these processes. Thus, the present method can provide He flux data which are free from the effects of elemental fractionation of CH_4 and He, and independent of the diffusion mechanism of He in the sediment.

Estimated ^3He and ^4He fluxes at the Chinshui and Chuhuangkeng sites are (3.9 ± 1.1) and $(2.7 \pm 0.6) \times 10^6$, and (7.2 ± 3.0) and $(2.4 \pm 0.8) \times 10^6 \text{ atoms cm}^{-2} \text{ s}^{-1}$, respectively. Even though the U and Th contents of the sedimentary rock are 50% higher than the assumed value, the changes in flux are not significant. In addition, the variation of R_s by a factor of 3 does not, in practice, affect the fluxes.

The estimated ^4He fluxes are ~ 30 times higher than that of Kilauea volcano, Hawaii⁵, and ~ 10 times that seen in fault zones in the Soviet Union⁷. O'Nions and Oxburgh⁶ reported an

^4He flux from the continents of $2.8 \times 10^6 \text{ atoms cm}^{-2} \text{ s}^{-1}$, based on the terrestrial heat flow and heat/He flux ratio. Their ^4He flux agrees well with the values of the present work. In the oceans, Craig *et al.*¹ reported a ^3He flux of 4 atoms cm^{-2} and a ^4He flux of $3 \times 10^5 \text{ atoms cm}^{-2} \text{ s}^{-1}$, based on the excess He in sea water. This ^3He value is equivalent to the present value, and their ^4He value is about one-tenth of ours. Direct measurements of the ^4He flux on the ocean floor^{3,4} give values of $0.11\text{--}1.6 \times 10^5 \text{ atoms cm}^{-2} \text{ s}^{-1}$, which are much lower than ours.

If it is plausible to take the present ^4He fluxes as representative of continental land areas (generalization of the ^4He fluxes may be supported by the agreement with the continental ^4He flux estimated by heat/He flux ratio⁶), there is a large discrepancy in ^4He fluxes between land and ocean areas. As the ^4He flux in land areas is significantly larger than that in the ocean, the major source of atmospheric ^4He may be attributed to continental crust in the land areas. The ^4He flux in the oceans contributes little to its atmospheric budget, even though the oceans dominate the surface of the Earth. This signature agrees with the recent estimation based on the $^3\text{He}/^4\text{He}$ ratios in pore waters of deep-sea sediments¹⁹.

We thank Mr G. Igarashi for providing the computer program for the nonlinear least-squares method.

Received 13 January; accepted 21 May 1986.

1. Craig, H., Clarke, W. B. & Beg, M. A. *Earth planet. Sci. Lett.* **26**, 125–132 (1975).
2. Bieri, R. H. & Koide, M. *J. geophys. Res.* **77**, 1667–1676 (1972).
3. Barnes, R. O. & Bieri, R. H. *Earth planet. Sci. Lett.* **28**, 331–336 (1976).
4. Sayles, F. L. & Jenkins, W. J. *Science* **214**, 13–18 (1982).
5. Naughton, J. J., Lee, J. H., Keeling, D., Finlayson, J. B. & Dority, G. *Science* **180**, 55–57 (1973).
6. O'Nions, R. K. & Oxburgh, E. R. *Nature* **306**, 429–431 (1983).
7. Tugarinov, A. I., Osipov, Yu. G. & Rentin, Yu. V. *Geokhimiya* **11**, 1616–1625 (1975).
8. Urabe, A., Tominaga, T., Nakamura, Y. & Wakita, H. *Geochem. J.* **19**, 11–25 (1985).
9. Sano, Y., Tominaga, T., Nakamura, Y. & Wakita, H. *Geochem. J.* **16**, 337–345 (1982).
10. Sano, Y. & Wakita, H. *J. geophys. Res.* **90**, 8729–8741 (1985).
11. Hunt, J. M. *Petroleum Geochemistry and Geology* 617 (Freeman, San Francisco, 1979).
12. Aldrich, L. T. & Nier, A. O. *Phys. Rev.* **74**, 1590–1594 (1948).
13. Tolstikhin, I. N. in *Terrestrial Rare Gases* (eds Alexander, E. C. & Ozima, M.) 33–62 (Japan Scientific Society Press, Tokyo, 1978).
14. Wakita, H. & Sano, Y. *Nature* **305**, 792–794 (1983).
15. Sato, K. *et al. Petrol. Geol. Taiwan* **7**, 281–293 (1970).
16. Mason, B. & Moore, C. B. *Principles of Geochemistry* 350 (Wiley, New York, 1982).
17. Morrison, P. & Pine, J. *Ann. N.Y. Acad. Sci.* **62**, 71–79 (1955).
18. Gerling, E. K., Mamyrin, B. A., Tolstikhin, I. N. & Yakovleva, S. S. *Geokhimiya* **5**, 608–617 (1971).
19. Sano, Y. *Geochem. J.* (in the press).

A brood parasitic catfish of mouthbrooding cichlid fishes in Lake Tanganyika

Tetsu Sato*

Department of Zoology, Faculty of Science, Kyoto University, Kitashirakawa-oiwakecho, Sakyo-ku, Kyoto 606, Japan

Brood parasitism, where a brood of the parasitic species is fostered by the parents of another species, is well known among birds¹. In most cases, such offspring show a complete reliance upon their host parents for food, protection and warmth until their independence. In other vertebrate groups, however, such total dependence upon a host species is unknown. I report here the first example of true brood parasitic behaviour discovered among fishes. In Lake Tanganyika, an endemic mochokid catfish, *Synodontis multipunctatus* Boulenger, is a brood parasite of mouthbrooding fishes of the family Cichlidae. The eggs of the catfish are incubated in the mouths of any of several host species together with the host's eggs, but hatch earlier. Following absorption of their yolk sacs, the catfish fry feed upon the fry of the host while still in its mouth. Thus the early stages of development of this catfish not only depend upon their hosts for food and protection, but exploit almost their entire parental investment.

Mouthbrooding of cichlids is one of the most advanced parental care systems known among fishes². The eggs of the mouthbrooder are usually picked up by the female immediately after oviposition and are incubated in the buccal cavity. Following yolk sac absorption, the fry frequently swim out to forage and return, using the mouth as a refuge until they become independent^{3,4}.

During a study of the parental care strategies of several Tanganyikan cichlid mouthbrooders at Uvira, Zaire, from July to December 1985, I collected 32 brooding female cichlids which contained eggs or fry of a mochokid catfish, *S. multipunctatus*, in their mouths (Fig. 1). These females represented six species (Table 1). Although an intensive search was made for the eggs or fry of *S. multipunctatus* in rock interstices and among piles of rocks in an area of ~300 m² where the brooding females were collected, they were found only in the mouths of the cichlid hosts. *S. multipunctatus* may therefore be regarded as an obligate brood parasite of mouthbrooders at least in this locality. Among the eight cases involving eggs of *S. multipunctatus*, seven were

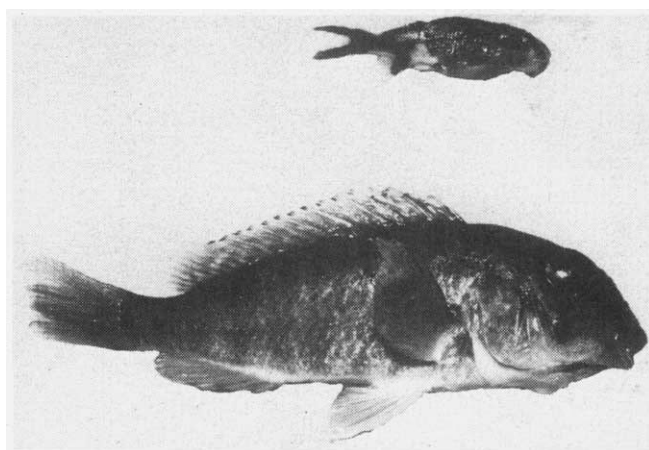


Fig. 1 A brooding female of *Simochromis babaulti* and a young *Synodontis multipunctatus* found in its mouth. No host fry were present. The catfish fry is the largest specimen (26.85 mm total length) collected from the mouth of any host. The standard length of the host is 58.65 mm.

associated with eggs of the host in the cleavage phase, whereas one brood of catfish eggs in which hatching was imminent, and two broods of catfish prolarvae, were associated with host eggs in the embryonic phase. Analysis of one host species, *Simochromis diagramma*, revealed the mean sizes of the catfish fry and associated host fry to be closely correlated ($r = 0.963$, $P < 0.001$). The catfish fry were never found with the host eggs and later invasion of the host by free-swimming catfish fry is unlikely. Therefore, although not observed, the eggs of the catfish are probably deposited somehow during the spawning bouts of the host.

Nineteen *S. multipunctatus* fry and 7 eggs from 10 broods were reared in aquaria. The larval yolk sac is absorbed within 3 days after hatch. If the fry of the host are present at this stage, the catfish fry begin to feed on them. Smaller fry bite the body of the host fry and gradually suck up the whole of their yolk whereas larger fry swallow the host fry whole. Stomach content analysis of wild-caught catfish fry (Table 2) revealed them to feed on host fry, and, on exhausting this food source, to resort to feeding on benthic materials and return. The mean number of host fry found associated with the catfish fry was significantly fewer than that found associated with eggs or prolarvae (Student's *t*-test, $P < 0.001$).

Incidental field data⁵ and some fragmentary aquarium observations^{6,7} have led to speculation on the possibility of brood parasitic behaviour in this species⁸. However, the present report is the first direct evidence based on field observations.

* Address for correspondence: c/o Professor H. Okumura, Department of Biology, Japan Women's University, 2-8-1 Mejirodai, Bunkyo-ku, Tokyo 112, Japan.

Table 1 The occurrence and number of *Synodontis multipunctatus* eggs and fry in the broods of six species of mouthbrooding cichlids

Species	No. of broods examined	No. of broods with catfish	% With catfish	Total no. of catfish	Mean no. in a brood
<i>Simochromis diagramma</i>	284	22 (6, 16)	7.7	65	3.0 (0.23)
<i>Simochromis babaulti</i>	89	3 (0, 3)	3.4	4	1.3 (0.04)
<i>Tropheus moorei</i>	35	2 (1, 1)	5.7	2	1.0 (0.06)
<i>Ctenochromis horei</i>	20	3 (1, 2)	15.0	19	6.3 (0.95)
<i>Gnathochromis pfefferi</i>	72	1 (0, 1)	1.4	2	2.0 (0.03)
<i>Pseudosimochromis curvifrons</i>	12	1 (0, 1)	8.3	1	1.0 (0.08)
Total	512	32 (8, 24)	6.3	93	2.9 (0.18)

Among 32 broods with *S. multipunctatus*, 24 broods contained catfish eggs or fry, together with 1–50 host eggs or fry, while 8 broods contained catfish fry only. The number of catfish eggs or fry in a brood ranged from 1 to 8. In the third column, numbers in parentheses represent the number of broods with *S. multipunctatus* eggs and fry respectively. Numbers in the last column represent the mean number of catfish per parasitized brood, and numbers in parentheses are mean number of catfish per brood examined. Thus, in total, one brood of host species contains a mean of 0.18 young of the catfish, that is, approximately 1 in 5 females of host species rear 1 catfish young. All samples were collected by nets from the rocky shore area at a depth of 0.2–1.8 m, where adults of *S. multipunctatus* were frequently observed. Samples were stored in 10% formalin and the number and sizes of eggs, fry and parents were measured in the laboratory. Care was taken to collect all the eggs and fry in the mouths.

Table 2 Results of stomach content analysis of *S. multipunctatus* fry collected from the mouths of hosts

	Size range (total length; mm)	No. of young which fed on:		
		Yolk-like material	Benthic material	Both
Together with host fry	10.83–23.45	9	2	0
Without host fry	16.03–26.85	1	5	1

Benthic material includes sand, ostracods and insects. Two specimens had fed on benthic material even though host fry were present in the mouths. In these cases, the host fry were the largest fry found associated with catfish fry ($15.90 \text{ mm} \pm 0.60 \text{ s.d.}$, $n = 10$ and $13.96 \text{ mm} \pm 0.20 \text{ s.d.}$, $n = 8$ in total length, respectively). However, whenever host fry were $< 13.5 \text{ mm}$, the catfish fry had always fed on yolk-like materials. Therefore, host fry above this size seem able to avoid predation by catfish fry. Two other specimens contained yolk-like materials in the digestive tracts although no host fry were present. They had probably consumed all the host fry shortly before collection.

Brood parasitism or interspecific brood care has been reported in several species of freshwater fishes^{9–14}, but such cases have usually been regarded as mutually altruistic behaviours or due to facultative mixing of free-swimming fry. Parasitism by *S. multipunctatus* is clearly disadvantageous to the hosts, and the mixing of broods must have been accomplished by the catfish parents during the spawning of the host. Therefore, it appears to be the only example of true brood parasitism recorded among fishes.

The development of the catfish eggs and fry and behaviour of the fry are apparently well adapted to exploit the host fry as a food source and the host parent as a refuge. These features suggest a long-established association between the catfish and mouthbrooders in the lake. Under normal circumstances, brooding female cichlids are exceptionally reliable parents. In one host species, *S. diagramma*, mortality of the young while under parental care is almost negligible (T.S., unpublished). Paradoxically, such a highly sophisticated parental care system has provided a niche for a specialized brood parasite.

I thank the staffs of CRSN/Uvira, Republique du Zaïre, for their collaboration and permission to undertake this research; and A. Rossiter, T. Abe, H. Kawanabe, R. J. McKaye and M. Yuma for reviewing the manuscript. This study was supported by the Grant-in-Aid for Overseas Scientific Survey 60041036 and 61043031 from the Ministry of Education, Science and Culture, Japan.

Received 15 April; accepted 27 May 1986.

1. Payne, R. B. *A. Rev. Ecol. Syst.* **8**, 1–28 (1977).
2. Keenleyside, M. H. A. *Diversity and Adaptation in Fish Behaviour* (Springer, Berlin, 1979).
3. Baerends, G. P. & Baerends-Van Roon, J. M. *Behaviour Suppl.* **1** (1950).
4. Fryer, G. & Iles, T. D. *The Cichlid Fishes of the Great Lakes of Africa* (Oliver & Boyd, Edinburgh, 1972).
5. Brichard, P. *Buntbar. Bull.* **74**, 10–12 (1979).
6. Ferguson, J. *Buntbar. Bull.* **98**, 13–16 (1983).
7. Finley, L. *Buntbar. Bull.* **98**, 17–18 (1983).
8. Finley, L. *Freshwater and Marine Aquarium* **7**(6), 22–25, 63, 66–75 (1984).
9. Goff, G. P. *Copeia* **1984**, 149–152 (1984).
10. McKaye, K. R. *Am. Nat.* **111**, 301–315 (1977).
11. McKaye, K. R. in *Natural Selection and Social Behavior* (eds Alexander, R. & Tinkle, D.) 173–183 (Chiron, New York, 1981).
12. McKaye, K. R. *Oecologia* **66**, 358–363 (1985).
13. Ribbink, A. J. *Nature* **267**, 243–244 (1977).
14. Ribbink, A. J., Marsh, A. C., Marsh, B. & Sharp, B. J. S. *Afr. J. Zool.* **15**, 1–6 (1980).

Novel sex differences in linkage values and meiotic chromosome behaviour in a marsupial

J. H. Bennett, D. L. Hayman & R. M. Hope

Department of Genetics, University of Adelaide, Adelaide, South Australia 5001

There have been few reported family studies of gene segregation in marsupials and no report on genetic linkage in a marsupial species. The paucity of such genetic data probably stems from the lack of marsupial species sufficiently well adapted to a laboratory environment to give the abundant progeny from controlled crosses which are required in such work. A colony of the fat-tailed insectivore *Sminthopsis crassicaudata* (Marsupialia: Dasyuridae), which was established in the R. A. Fisher Laboratories, University of Adelaide, more than 20 years ago¹, has now provided unusual linkage data. These data indicate that the linkage situation in this marsupial species differs markedly from that in eutherian mammals. Extremely large differences exist between the sexes in the values of the recombination frequencies, with much closer linkage occurring in females. Cytological examination of meiosis in males and females has revealed a major difference between the sexes in chromosome behaviour.

The primary aim in setting up a colony of *S. crassicaudata* was to develop a small laboratory-bred marsupial suitable for intensive genetic studies and showing as much genetic variation as possible. Most of the foundation animals were taken from the wild in scattered parts of central and southern Australia. Computerized records are maintained on the parents of all individuals bred in the colony and these make possible rapid determination of the ancestry of any animal. The ancestry of

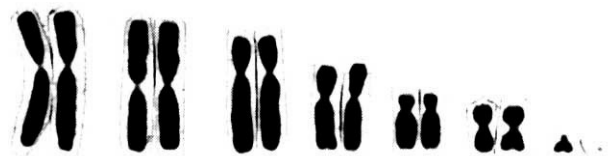


Fig. 1 Karyotype of male *Sminthopsis crassicaudata*.

the present colony of about 800 individuals traces back to about 30 foundation animals introduced before 1975. Since 1975, the colony has been maintained as a closed self-sustaining population. Under natural conditions, births are said to occur in the period from July to February², but with the lighting and temperature regime that we use, they occur almost uniformly throughout the year. Adult animals weigh about 14 g and the young are born at a very immature stage, their weight being about 1,000th that of an adult. A maximum of 10 litter young can be reared in a pouch. Weaning occurs when the litter is about 65 days old and although some litters of 10 have been reared to weaning, the average litter size at weaning is about 5. As the mother usually enters oestrus shortly after suckling ceases, and the gestation period is about 13 days, another litter may be born about 80 days after the earlier one. Females come into oestrus first when about 4 months old and some have produced 4 litters in the following 12 months.

Genetic variants for a number of characters are segregating in the colony. Analysis of blood samples using electrophoresis and isoelectric focusing has yielded extensive family data on variants of the following seven proteins: adenosine deaminase (ADA; EC 3.5.4.4), amylase (AMY; EC 3.2.1.1), glucose phosphate isomerase (GPI; EC 5.3.1.9), 6-phosphogluconate dehydrogenase (PGD; EC 1.1.1.44), protease inhibitor (PI), super-

Table 1 Linkage data from double test-cross progenies of *S. crassicaudata* shown separately for female and male double heterozygotes

Gene loci	Female			Male		
	No. of recombinants	Total no. of progeny	r.f.	No. of recombinants	Total no. of progeny	r.f.
Linkage group I						
ADA-GPI	0	211	0	88	236	0.37
ADA-PI	0	61	0	26	78	0.33
GPI-PI	0	11	0	29	116	0.25
Linkage group II						
SOD-TRF	0	169	0	55	335	0.16
SOD-PGD	6	97	0.06	99	197	0.50
PGD-TRF	52	419	0.12	172	334	0.51

r.f., Recombination fraction.

oxide dismutase (SOD; EC 1.15.1.1) and transferrin (TRF). For each of these proteins, the variants show autosomal inheritance, the family data being in close accord with mendelian expectations³⁻⁶. Data obtained from double test-cross matings show that three of these genetic markers (ADA, GPI, PI) are in one linkage group and another three (PGD, SOD, TRF) are in a second linkage group. The remaining marker (AMY) shows independent assortment with all of the others. The linkage values obtained from double test-cross progenies for pairs of these linked markers are shown in Table 1. These are the first linkage data to be reported for a marsupial species and they are remarkable for the extremely large differences between the values of the recombination fraction in males and females, with much closer linkage occurring in females. For each of the three pairs of loci in linkage group I, no genetic recombination has yet been observed in any test-cross progeny when the doubly heterozygous parent is female, whereas when the doubly heterozygous parent is a male many recombinants are found. The same situation is also found for one pair of loci (SOD and TRF) in linkage group II. The other two pairs of loci in linkage group II (PGD and SOD, PGD and TRF) provide the only examples yet available in which recombination of linked genes has been observed in the progeny of doubly heterozygous females, and it is notable that doubly heterozygous males show independent assortment for both of these two pairs of gene loci.

S. crassicaudata has six pairs of autosomal chromosomes and a pair of sex chromosomes (Fig. 1). At diakinesis and metaphase I of meiosis in the male, the autosomes have an average of 13.6 chiasmata per cell (based on 20 cells) and more than half of these are clearly interstitial (Fig. 2). At a comparable stage in the female there are, on average, 10.2 chiasmata (based on 5 entire cells where all chromosomes are separate) and all of the chiasmata seen in a total of 11 cells are terminal or sub-terminal. These observations provide a physical parallel to the observed difference in genetic recombination between the sexes.

There has been no systematic study of meiosis in the female of any other marsupial species and so it is not known whether the difference in meiotic chromosome behaviour observed in *S. crassicaudata* is a consistent feature of all marsupials. However, isolated observations of meiosis in males and females in some species of the *Phalangerioidea* suggest that this is not the case⁷.

Since distal localization of chiasmata in females is observed for all of the autosomes in *S. crassicaudata*, the observed sexual disparity in recombination is expected to be a general characteristic of all of the autosomal linkage groups. The small size of the sex chromosome bivalent precludes determination of chiasma localization. A search for X-linked gene markers to be used in a study of recombination has been unsuccessful.

In contrast to our findings, the sex differences reported on intra-chromosomal recombination in eutherian mammals (for example, in man and mouse) are much less extreme and males rather than females generally show closer linkage. The evidence from eutherian species does not support any consistent rule relating a marked reduction in recombination frequency to

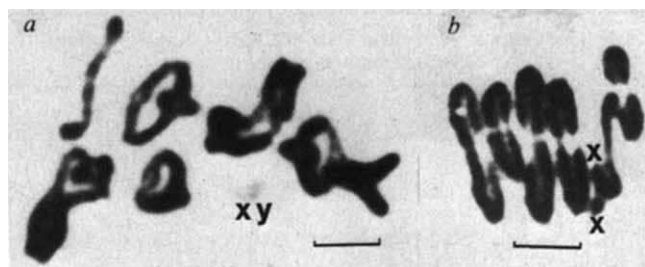


Fig. 2 Metaphase I of meiosis in *S. crassicaudata*. *a*, Male; many interstitial chiasmata are apparent. XY, indicates faintly staining sex chromosome bivalent. *b*, Female; chiasmata are terminally localized. XX, indicates the X bivalent. Ova dissected out from animals injected with 10 IU pregnant mare serum (48 h) and 5 IU human chorionic gonadotropin and prepared after Tarkowski¹¹. Scale bar, 5 μ m.

heterogamety⁸, contrary to the early suggestion by Haldane⁹ and Huxley¹⁰; certainly, our observations are in sharp conflict with their suggestion.

We thank Clive Chesson and Mrs D. Golding for technical assistance and Dr P. J. Quinn for technical advice. This work was supported by the Australian Research Grants Scheme and a University of Adelaide research grant.

Received 9 April; accepted 4 June 1986.

1. Bennett, J. H., Smith, M. J., Hope, R. M. & Chesson, C. M. in *The Management of Australian Mammals in Captivity* (ed. Evans, D. D.) 38-44 (Zoological Board of Victoria, Melbourne, 1982).
2. Morton, S. R. *Aust. Wildl. Res.* **5**, 163-182 (1978).
3. Hope, R. M. & Godfrey, G. K. *Aust. J. Biol. Sci.* **21**, 587-591 (1968).
4. Cooper, D. W. & Hope, R. M. *Biochem. Genet.* **5**, 65-68 (1971).
5. Hope, R. M. *Anim. Blood Groups biochem. Genet.* **3**, 95-99 (1972).
6. Hope, R. M., Bennett, J. H., Chesson, C. M. & Adams, M. *Biochem. Genet.* **22**, 221-229 (1984).
7. Sharp, P. J. thesis, Univ. Adelaide (1984).
8. Dunn, L. C. & Bennett, D. *Genet. Res.* **9**, 211-220 (1967).
9. Haldane, J. B. S. *J. Genet.* **12**, 101-109 (1922).
10. Huxley, J. S. *J. Genet.* **20**, 145-156 (1928).
11. Tarkowski, A. K. *Cytogenetics* **5**, 394-400 (1966).

A molecular link between the bats of New Zealand and South America

Elizabeth D. Pierson*, Vincent M. Sarich†, Jerold M. Lowenstein‡, Mike J. Daniels§ & William E. Rainey*

* Museum of Vertebrate Zoology and Department of Zoology, University of California, Berkeley, California 94720, USA

† Departments of Anthropology and Biochemistry, University of California, Berkeley, California, 94720, USA

‡ Department of Medicine, University of California, San Francisco, California 94143, USA

§ Ecology Division, Department of Scientific and Industrial Research, Lower Hutt, New Zealand

Along with the kiwis (*Apteryx*), tuatara (*Sphenodon*) and leiopelmatid frogs, the now rare lesser short-tailed bat (*Mystacina tuberculata*), one of only two species^{1,2} in the endemic family *Mystacinidae*, has long been viewed as one of New Zealand's archaic, mystery vertebrates, and has presented taxonomists with a major puzzle since its first description in 1843 (ref. 3). We report here the results of immunological comparisons involving the albumin and transferrin of *Mystacina* which indicate that its closest phylogenetic affinities are with the New World phyllostomoids—noctilionids, mormoopids and phyllostomids. We estimate the separation between the *Noctilio* and *Mystacina* lineages to have occurred about 35 Myr ago.

Fig. 1 The most widely accepted classification of bats, with a division into 4 superfamilies and 16 families¹¹. The asterisk indicates families with which *Mystacina* has been associated.

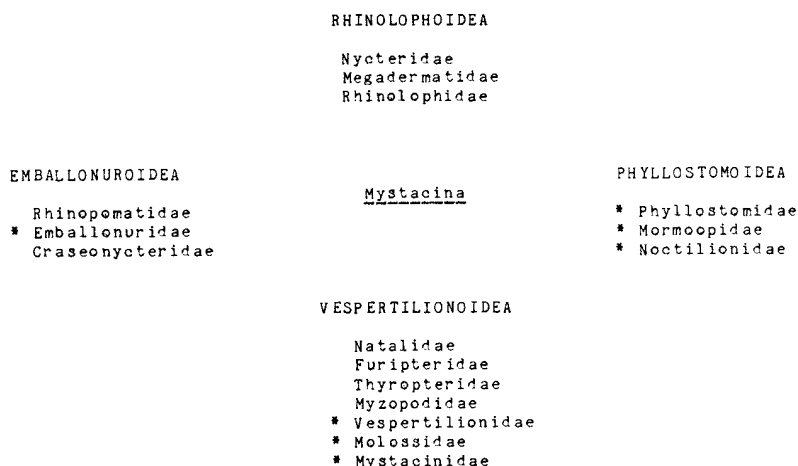
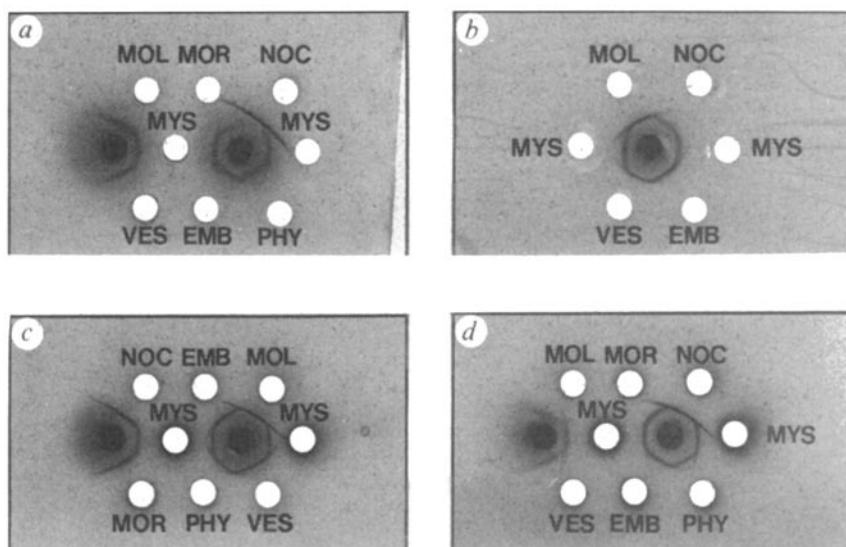


Fig. 2 Immunodiffusion gels using albumin and transferrin antisera. *a*, Anti-*Noctilio leporinus* albumin; *b*, anti-*Cabassous/Tamandua/Bradypus/Felis/Canis/Hyene/Ursus/Aotus/Lemur* albumins (pooled); *c*, anti-*Tadarida brasiliensis* transferrin; *d*, anti-*Noctilio albiventris* transferrin. These immunodiffusion gels were prepared using 1.2% DifcoBacto agar in isotris²², 4.0% polyethylene glycol (relative molecular mass, 8,000), and stained with Serva Blue G. Antiserum (10 µl) was loaded in the dark central wells, and plasma dilutions (10 µl; 1:15 for transferrins, 1:50 for albumins) in the surrounding, lighter wells. The precipitin lines, forming from the reaction of the antigens with the antisera, can be interpreted phylogenetically. Fusion of the precipitin lines indicates immunological equivalence (that is, taxa are at similar immunological distances from the antiserum taxon); spurring of one line over another indicates that the spurring taxon is immunologically closer to the antiserum taxon. The antigens were from the following taxa: MOL, *Tadarida brasiliensis* (Molossidae); MOR, *Pteronotus parnellii* (Mormoopidae); NOC, *Noctilio albiventris* (Noctilionidae); VES, *Antrozous pallidus* (Vespertilionidae); EMB, *Saccopteryx bilineata* (Emballonuridae); PHY, *Vampyrus spectrum* (Phyllostomidae); MYS, *Mystacina tuberculata* (Mystacinidae).



Mystacina tuberculata, although a competent flier, is, like New Zealand's many flightless birds, markedly terrestrial. Moreover, showing ecological versatility unique for a bat, it feeds on aerial and ground insects, fruit, pollen, nectar, and possibly small vertebrates (nestling birds) and carrion^{4,5}. Not surprisingly, it shows a striking complex of morphological adaptations: extremely robust hindlimbs; velvety, mole-like fur; a unique basal talon on each claw; sturdy chisel-like incisors; an unusually extensible, papillated tongue; and wings that fold into a skin pocket beneath a protective, leathery sheath^{4,6,7}.

This assemblage of distinctive features strongly suggests lengthy *in situ* evolution, and has, quite predictably, led to taxonomic confusion. *Mystacina* has, at one time or another, been placed in 6 of the 16 recognized bat families and in three of the four microchiropteran superfamilies^{4,8-16} (Fig. 1). Since 1907¹², it has been accorded separate family status (Mystacinidae), and since 1945¹³, has been viewed as an aberrant member of the superfamily Vespertilionoidea. This latter placement remains largely unsubstantiated, and has been questioned by two New Zealand scientists (including M.J.D.), who have worked closely with this bat (ref. 4 and P. D. Dwyer, unpublished data).

Problems of this sort have yielded to comparative molecular

approaches, as, for example, the demonstration of monophyly in ratite birds^{17,18} and the placement of the giant panda among the bears¹⁹. Nonetheless, the placement of *Mystacina* was not straightforward. We began by conducting a general survey, reacting *Mystacina* albumin with antisera to the albumins of numerous bat species by standard Ouchterlony immunodiffusion²⁰. It uniformly gave very weak reactions, indicating that it had experienced a greater than average amount of change. This immediately focused our attention on the phyllostomoids of Central and South America, the only other group of bats with comparably divergent albumins²¹. Further quantitation was difficult to achieve. The large immunological distances involved meant that the albumin in complement itself (guinea pig plasma) would interfere with any microcomplement fixation (MC/F) reactions²²; lack of material precluded the alternative of quantitative precipitin analysis²³. We then turned to solid-phase radioimmunoassay (RIA)²⁴⁻²⁶, which is not subject to either of these limitations and has high resolving power at large immunological distances. A RIA survey, using representative albumins and transferrins of all major bat lineages, strongly suggested a phyllostomoid association, indicating a special relationship between *Mystacina* and the bulldog bats, *Noctilio* (Table 1). We tested this, first by a modification of the Ouchter-

Table 1 Immunological similarity of *Mystacina* to other bats

	RIA*		Noc.	Mor.	MCF anti-transferrin			Bal.	Mac†.
					Glo.	Ant.	Tad.		
<i>Noctilio</i>	0.88	<i>Noctilio</i>	—	97	99	134	110	127	151
<i>Phyllostomus</i>	0.59	<i>Mystacina</i>	105	104	125	136	137	131	155
<i>Pteronotus</i>	0.33								
<i>Myotis</i>	0.51	<i>Mormoops</i>	90						
<i>Tadarida</i>	0.50	<i>Pteronotus</i>	100						
<i>Balantiopteryx</i>	0.05	<i>Glossophaga</i>	121						
<i>Pteropus</i>	0.01	<i>Antrozous</i>	146						
		<i>Tadarida</i>	155						
		<i>Balantiopteryx</i>	132						
		<i>Macroderma</i>	172						

The taxa used in these comparisons represent all the families with which *Mystacina* has been associated: *Noctilio albigentris* (Noc) (Noctilionidae); *Mormoops megalophylla* (Mor) and *Pteronotus parnellii* (Phyllostomidae); *Antrozous pallidus* (Ant) and *Myotis yumanensis* (Vespertilionidae); *Tadarida brasiliensis* (Tad) (Molossidae); *Balantiopteryx plicata* (Bal) (Emballonuridae).

* Antisera to the albumins and transferrins of the bats named here were reacted in solid phase radioimmunoassay with the serum of each of these genera and of *Mystacina*. The numbers are correlation coefficients (means for albumin and transferrin) of the set of *Mystacina* reactions compared with those of each of the other bats. A correlation coefficient of 1.0 would indicate identical reactions; -1.0 is the greatest possible dissimilarity. Based on this analysis, *Mystacina* is immunologically much more similar to *Noctilio* than to any other bat tested.

† The immunological distances from an outgroup, *Macroderma gigas* (Megadermatidae), indicate comparable amounts of change for *Noctilio* and *Mystacina* transferrins, showing that the relatively small immunological distance between *Mystacina* and *Noctilio* does not result from a lack of change along the *Mystacina* lineage.

Table 2 Derived characters uniting the phyllostomoid clade

Taxa	Upper incisors Broad	Upper incisors contiguous	Clitoris elongate	Vulval opening longitudinal	Os penis present
Mystacinidae	+	+	+	—	(+)
Phyllostomidae	+	+	+	+	—
Mormoopidae					
<i>Mormoops</i>	+	+	+	+	(+)
<i>Pteronotus</i>	+	+	+	+	—
Noctilionidae	—	+	+	+	—
Most other extant bat families	—	—	—	—	+
Icaronycteridae*	—	—	NI	NI	+

+, Condition present; —, Condition absent; (+), residual condition; NI, no information.

* Eocene fossil³⁶.

lony procedure, and then by MCF comparisons of transferrins.

The addition of polyethylene glycol (an agent known to enhance protein precipitation)²⁷ to immunodiffusion gels was found to greatly expand the phylogenetic scope of heterologous reactions (W.E.R., unpublished). This property made it possible to carry out a broad comparative survey involving both the albumin and transferrin of *Mystacina*. The results (Fig. 2) are unequivocal in associating *Mystacina* with the phyllostomoids. We were especially impressed that, even though *Noctilio* and *Mystacina* have the most divergent of all bat albumins tested (Fig. 2b), nonetheless, with anti-*Noctilio* albumin, *Mystacina* albumin spurs over that of every other bat tested (Fig. 2a). The transferrin results (Fig. 2c, d) corroborate the albumin findings: with antisera to non-phyllostomoid transferrins, *Mystacina* transferrin fails to spur over that of any non-phyllostomoid; with anti-*Noctilio* transferrin, however, *Mystacina* reacts better than any other taxon (except *Pteronotus*), including the phyllostomids.

In the MCF transferrin comparisons (Table 1), *Mystacina* reacts best with two phyllostomoids, *Noctilio* and *Mormoops*, and shows no affinities with any non-phyllostomoid. Since branching order among phyllostomoid taxa is still unclear, we defer more precise placement of *Mystacina* until antisera to its transferrin and albumin are available.

The discovery that *Mystacina* is, at least in part, a fruit- and nectar-feeding bat^{4,28} provided the first recent clue to its phyllostomoid affinities (frugivory and nectar-feeding being major adaptations in only two bat families, the pteropids and phyllo-

stomids), but this proposed association was, in fact, originally suggested by Gray, who first described *Mystacina* in 1843³, and later, in 1863, by Tomes¹⁵. That a New Zealand bat shows affinities with tropical American taxa should not be surprising. Even though New Zealand has been isolated for some 80 million years²⁹, much of its flora and fauna show South American affinities.

If we re-examine those morphological features which have been used in the past to classify *Mystacina*, with vespertilionids, emballonurids and molossids, all would seem to be either primitive (for example, ear shape) or convergent (for example, strong hindlimbs). In contrast, we have identified five derived morphological features which characterize the phyllostomoid clade³⁰⁻³² (Table 2), and for three of the five *Mystacina* shares the derived phyllostomoid condition.

Although we are not as yet entirely comfortable with a molecular time scale for bat evolution, and the highly divergent nature of the phyllostomoid albumin evolution is a complicating factor, our transferrin data do suggest mean maximum within bat immunological distances (IDs) to be about 10% less than those found within primates³³ and 10% greater than those found within carnivores³⁴. If we accept a Palaeocene origin for bats, this suggests a transferrin rate of about 3 ID units per Myr, indicating about 35 Myr of separation between *Noctilio* and *Mystacina*.

Given this time-frame, and the highly derived *Mystacina* morphology, it is tempting to speculate as to why there are not more *Mystacina*-derived lineages in New Zealand. After all,

avian lineages have produced many indigenous forms there. The answer may be that the phyllostomoids, for all their diversity in Central and South America, are an essentially tropical radiation, with very few species ever venturing north of Mexico. No phyllostomid is known from a climate as temperate as that of New Zealand. Thus, the lack of diversity may simply reflect the limitations of phyllostomid physiology, and we may wonder that *Mystacina* survived at all through the recent New Zealand glaciations³⁵.

We thank the governments of Australia, Costa Rica, New Zealand and Trinidad for permission to collect and export samples; R. J. Baker, R. L. Honeycutt, and the Texas Tech Museum for additional samples; K. F. Koopman, M. J. Novacek and J. L. Patton for editorial comments.

Received 21 March; accepted 19 June 1986.

1. Daniel, M. J. & Williams, G. R. *N.Z. J. Ecol.* **7**, 9–25 (1984).
2. Hill, J. E. & Daniel, M. J. *Bull. Br. Mus. nat. Hist.* **D48**, 279–300 (1985).
3. Gray, J. E. in *Travels in New Zealand* Vol. 2 (ed. Dieffenbach, E.) 181–182 (Murray, London, 1843).
4. Daniel, M. J. *N.Z. J. Zool.* **6**, 357–370 (1979).
5. Stead, E. F. *Trans. R. Soc. N.Z.* **66**, 188–191 (1937).
6. Dobson, G. E. *Proc. zool. Soc. Lond.* 486–488 (1876).
7. Dwyer, P. D. *Zool. Publ. Vict. Univ. Wellington* **28**, 1–28 (1962).
8. Dobson, G. E. *Ann. Mag. nat. Hist.* **16**, Ser. 4, 345–357 (1875).
9. Gray, J. E. *Ann. Mag. nat. Hist.* **17**, Ser. 3, 89–93 (1866).
10. Hutton, F. W. *Trans. Proc. N.Z. Inst.* **4**, 184–186 (1872).
11. Koopman, K. F. *Bat. Res. News* **24**, 36–38 (1983).

12. Miller, G. S. *Bull. U.S. natn. Mus.* **57**, 1–261 (1907).
13. Simpson, G. G. *Bull. Am. Mus. nat. Hist.* **85**, 1–350 (1945).
14. Tomes, R. F. *Proc. zool. Soc. Lond.* 134–135 (1857).
15. Tomes, R. F. *Proc. zool. Soc. Lond.* 81–84 (1863).
16. Van Valen, L. *Evol. Theory* **4**, 103–121 (1979).
17. Prager, E. M., Wilson, A. C., Osuga, D. T. & Feeney, R. E. *J. molec. Evol.* **8**, 283–294 (1976).
18. Sibley, C. G. & Ahlquist, J. E. in *Evolution Today* (eds Scudder, G. G. E. & Reveal, J. L.) 301–335 (Carnegie Mellon University, Pittsburgh, 1981).
19. Sarich, V. M. *Nature* **245**, 218–220 (1973).
20. Nilsson, L.-A. *Scand. J. Immun.* **17**, Suppl. 10, 57–68 (1983).
21. Arnold, M. L., Honeycutt, R. L., Baker, R. J., Sarich, V. M. & Jones, Jr., J. K. *Occas. Papers Mus. Texas Tech.* **77**, 1–15 (1982).
22. Champion, A. B., Prager, E. M., Wachter, D. & Wilson, A. C. in *Biochemical and Immunological Taxonomy of Animals* (ed. Wright, C. A.) 397–416 (Academic, New York, 1971).
23. Li, I. W., & Williams, C. A. in *Methods in Immunology and Immunochemistry* Vol. 3 (eds Williams, C. A. & Chase, M. W.) 94–102 (Academic, New York, 1974).
24. Lowenstein, J. M. *Naturwissenschaften* **67**, 343–346 (1980).
25. Lowenstein, J. M., Sarich, V. M. & Richardson, B. J. *Nature* **291**, 409–411 (1981).
26. Rainey, W. E., Lowenstein, J. M., Sarich, V. M. & Magor, D. M. *Naturwissenschaften* **71**, 586–588 (1984).
27. Kostner, G. & Holasek, A. *Analyt. Biochem.* **46**, 680–683 (1972).
28. Daniel, M. J. *N.Z. J. Zool.* **3**, 391–398 (1976).
29. Griffiths, J. R. & Varne, R. *Nature phys. Sci.* **235**, 83–86 (1972).
30. Smith, J. D. *Misc. Publs Mus. natn. Hist. Univ. Kans.* **56**, 1–132 (1972).
31. Luckett, W. P. in *Proc. 5th int. Bat. Res. Conf.* (eds Wilson, D. E. & Gardner, A. L.) 245–266 (Texas Tech Press, Lubbock, 1980).
32. Wood Jones, F. *J. Anat. Physiol.* **51**, 36–60 (1916–17).
33. Sarich, V. M. & Cronin, J. E. in *Molecular Anthropology* (eds Goodman, M. & Tashian, R. E.) 141–170 (Plenum, New York, 1976).
34. Sarich, V. M. *Trans. zool. Soc. Lond.* **33**, 165–171 (1976).
35. Fleming, C. A. *The Geological History of New Zealand and its Life* (Oxford University Press, 1979).
36. Jenson, G. L. in *Biology of Bats* Vol. 1 (ed. Wimsatt, W. W.) 1–62 (Academic, New York, 1970).

The mechanism of rectification at the electrotonic motor giant synapse of the crayfish

Stewart W. Jaslove & Peter R. Brink

Department of Anatomical Sciences, Health Sciences Center, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

The synapse between the giant interneurone and the motor giant axon of the crayfish is a well-known example of the rare class of current-rectifying electrotonic synapses^{1–3}. One early proposal for the basis of this rectification was that rectifying junctions are like diodes². Biological correlates of diodes can exist, such as constant-field channels which rectify by very high-speed rearrangements of charge carriers⁴, but these require high selectivity and large concentration gradients. Electrotonic synapses are believed to be composed of wide-bore (1–2 nm) gap-junction channels which have poor selectivity and bridge similar intracellular compartments³. An alternative mechanism for rectification would be by voltage-dependent gates that sense trans-synaptic potential. These two mechanisms can be distinguished because a diode should rectify instantaneously (on a biological time-scale) while a gated channel should show kinetic processes. Although a gating model is more consistent with the known behaviour of channels than a diode model, previous work has failed to find any time course for the rectification^{2,5–8}. We have now developed a high-quality voltage clamp and by working at reduced temperatures we are able to demonstrate channel kinetics. These results support the hypothesis that this rectifying synapse contains voltage-dependent gates.

Synapses were from the second and third abdominal ganglia of the ventral nerve cord of the crayfish *Procambarus clarkii* (Fig. 1). The instrumentation utilized a double voltage-clamp technique⁹, in which a pair of coupled axons are individually voltage-clamped with respect to bath ground, thereby clamping the junction between them. A fine glass cannula was inserted into a cut end of each of the two axons and advanced to the location of the synapse to provide a low-resistance path for current injection¹⁰. Each axon was also impaled in the synaptic region with a microelectrode for voltage measurement.

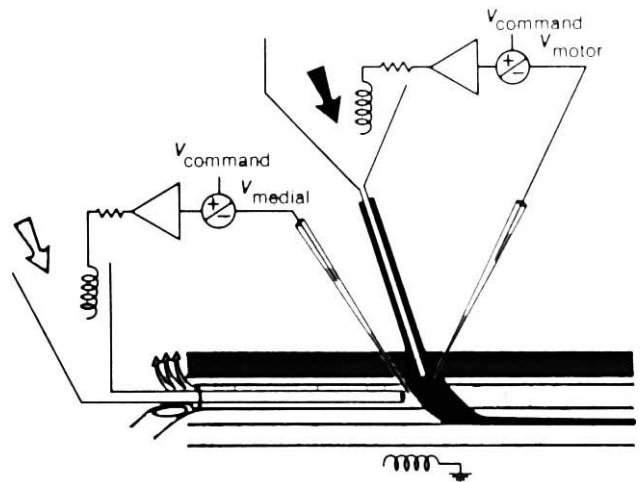


Fig. 1 Diagram showing the approximate relations of a single motor giant axon to its three presynaptic giant axons from a second or third abdominal ganglion. Motor axon, black; lateral axon, dark shading; medial axons, light shading. All four axons are typically 80–100 μm in diameter in the synaptic region. The cannulation shown is of a motor axon and an ipsilateral medial axon, which was the usual procedure for these experiments. The length constant of a medial giant axon is 2–4 mm¹⁶. Therefore, all electrode tips were positioned within 100 μm of the synaptic region to minimize axial voltage gradients. The cut end of the motor axon was 1–1.5 mm from the synapse and its cannula was 20 μm outer diameter; the cut end of the presynaptic axon was 2–3 mm from the synapse and its cannula was 40 μm o.d. Current loss to ground through the cut ends was minimized by the long length and reduced effective cross-sectional area of the cannulated axons. In any case, a current shunt at that point will not affect junctional current and voltage provided that total system current remains within the compliance of the current-passing circuitry. The use of cannulae also allowed for internal perfusion of the presynaptic giant fibres¹⁰. In general the motor axon was not well perfused, because it did not survive bulk flow well, but internal solution probably diffused out of the tip of its cannula and into its axoplasm during the course of an experiment (up to ~3 h). The composition of the internal solution was 220 mM CsF; 10 mM NaCl; 10 mM HEPES, pH 7.5. The external solution contained 210 mM NaCl; 15 mM CaCl_2 and 5 mM HEPES, pH 7.4. Membrane potassium currents were eliminated by internal Cs¹⁷ and sodium currents blocked by 300 nM external tetrodotoxin (TTX) to improve voltage-clamp control.

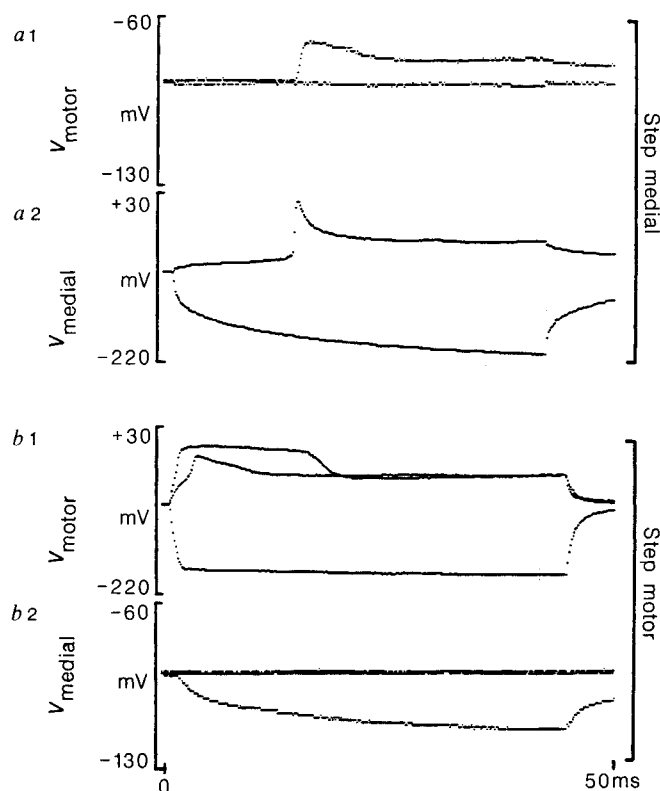


Fig. 2 Current-clamp records from a cannulated preparation, taken at 11.5°C. Both transmembrane potentials were set to -88 mV by the application of constant-bias currents through the cannulae. Constant current pulses were injected into the medial (*a2*) or motor (*b1*) axon and synaptic potentials were recorded in the trans-synaptic axon (*a1, b2*). The rectifying ability of this synapse is demonstrated by the failure of both the hyperpolarization of the medial axon to pass into the motor axon (*a1*) and the depolarization of the motor axon to pass into the medial axon (*b2*), whereas polarizations of opposite polarity are transferred readily. Two depolarizations of slightly different amplitude were applied to the motor axon (*b1*), to demonstrate the two kinds of active responses which it produces. The longer response has a higher threshold and is TTX-insensitive.

For the purpose of this paper, transmembrane potential (V_{medial} or V_{motor}) refers to the internal voltage of either axon with respect to bath ground and transjunctional potential (V_j) refers to the voltage across the junction ($V_{\text{medial}} - V_{\text{motor}}$). A positive voltage step across the junction represents either depolarization of the presynaptic giant fibre or hyperpolarization of the postsynaptic motor fibre.

Cannulated preparations gave current-clamp results comparable to those previously reported for intact, four-microelectrode systems^{2,5-7} (Fig. 2). Input resistances (R_i) of 19 well-rectifying preparations, measured by applying strong hyperpolarizing pulses, were $120 \pm 8 \text{ k}\Omega$ (mean $\pm$ s.e.m.) for the presynaptic axon and $93 \pm 7 \text{ k}\Omega$ for the motor axon. Assuming that the input resistance of the hyperpolarized motor axon is shunted by three motor synapses, and using the value for synaptic resistance (R_j) given below (345 k Ω), the actual input resistance of the motor axon would be as high as $\sim 230 \text{ k}\Omega$.

Recordings from another preparation under voltage-clamp control are depicted in Fig. 3. The difference in holding potentials ($\Delta \text{h.p.}$) was set to -30 mV and one or the other axon was stepped (Fig. 3*a2, b1*). The resulting transjunctional currents (I_j , Fig. 3*a1, b2*) were directly proportional to synaptic resistance⁹,

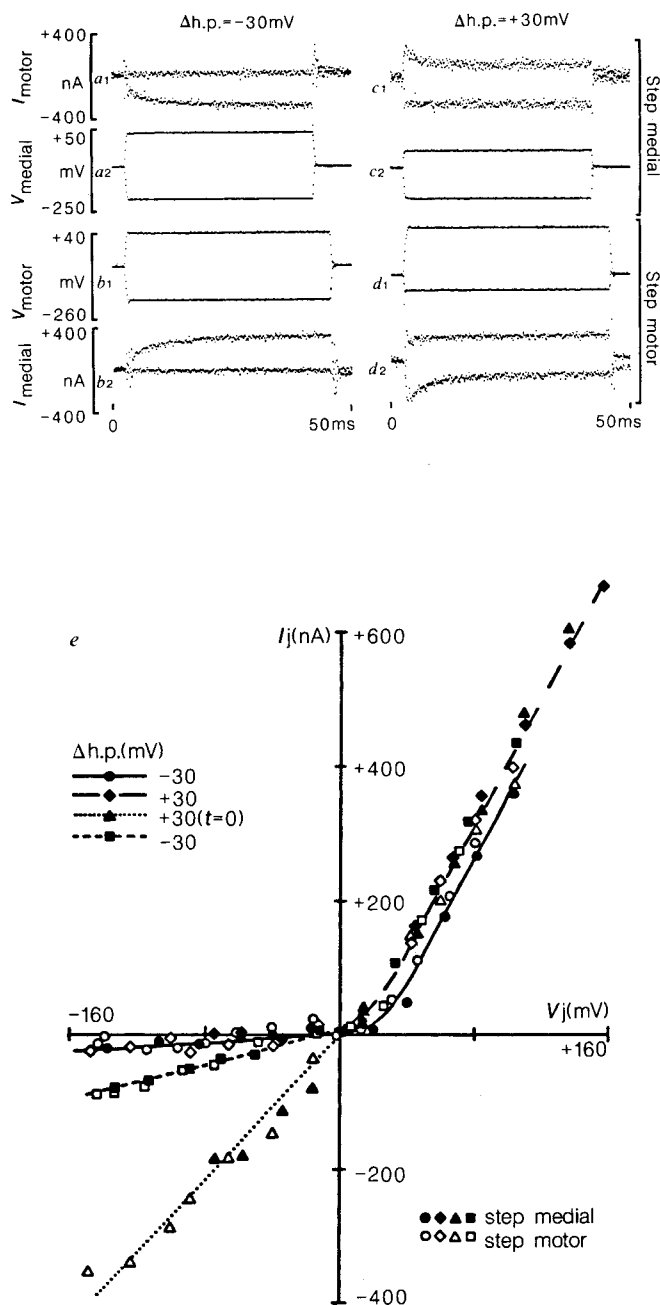


Fig. 3 Voltage-clamp data taken at 12°C. Transmembrane voltage (V_{medial} or V_{motor}) was stepped in one axon and transjunctional current (I_j) was recorded as I_{medial} or I_{motor} from the coupled axon. In interpreting these records, one must bear in mind that a polarization of one axon requires a current of opposite polarity to clamp the coupled axon. *a, b*, $\text{h.p.}_{\text{medial}} = -90 \text{ mV}$, $\text{h.p.}_{\text{motor}} = -60 \text{ mV}$, $\Delta \text{h.p.} = -30 \text{ mV}$; *c, d*, $\text{h.p.}_{\text{medial}} = -90 \text{ mV}$, $\text{h.p.}_{\text{motor}} = -120 \text{ mV}$, $\Delta \text{h.p.} = +30 \text{ mV}$. Note that steady-state current flows only in the medial to motor axon direction (small offsets were due to changes in the holding current). Slow kinetics represent channel activation (*a1, b2*) and deactivation (*c1, d2*). The apparent fast rising phase ($< 1 \text{ ms}$) on some of the I_j traces (*c1, d2*) is a reflection of the rise-time of the voltage clamp. *e*, Comparison of steady-state and instantaneous current-voltage relations of the complete experiment. Solid symbols represent data obtained by pulsing the medial axon and open symbols by pulsing the motor axon. Steady-state points ($\bullet, \blacklozenge, \blacksquare$) were recorded at 40 ms. Instantaneous points ($\blacktriangle$) were extrapolated to time $t = 0$ by approximating the early part of the decay in synaptic current as an exponential function. Positive current flow is defined as the direction from the medial to the motor axon. Smooth curves were drawn by eye.

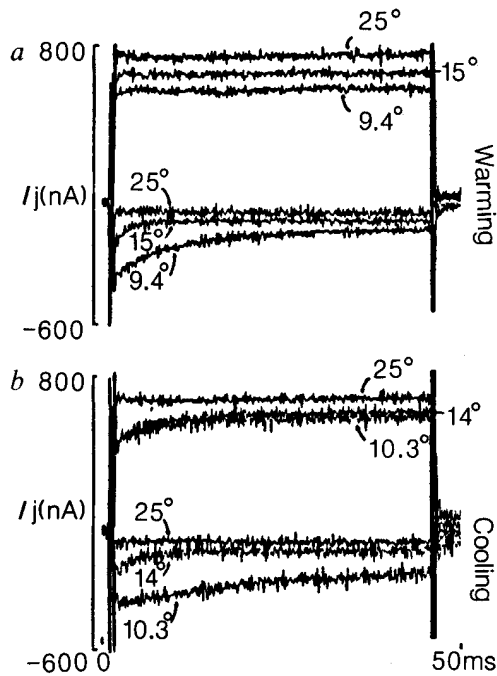


Fig. 4 Voltage-clamp records demonstrating the effect of temperature on synaptic gating. I_j was recorded from a medial axon while stepping the motor axon ± 85 mV and warming the preparation from 9.4°C to 25°C (a) and then cooling to 10.3°C (b). For reference, the time constant of decay at 9.4°C was about 7.5 ms. h.p._{medial} = -90 mV; h.p._{motor} = -90 mV; Δ h.p. = 0 mV. The recovery showed significant hysteresis in this experiment.

which reached a maximum value of 4.7 ± 1.2 M Ω for negative V_j s and a minimum value of 345 ± 61 k Ω for positive V_j s ($n = 7$).

To test whether the rising phase of I_j represented gating kinetics, Δ h.p. was changed to +30 mV to hold the junction in a conducting state. With pulses of either polarity (Fig. 3c2,d1) I_j now rapidly jumped to a high level and, at negative V_j s, quickly deactivated (Fig. 3c1,d2). These results indicate that at Δ h.p. = -30 mV most channels were closed and the rising phase of I_j was due to channel-opening kinetics, whereas at Δ h.p. = +30 mV most channels were open and the falling phase of I_j represented channel-closing kinetics.

The steady-state and instantaneous current-voltage relations from this experiment (Fig. 3e) raise several points. First, the insensitivity of I_j to the axon which is pulsed (medial or motor) suggests that R_j depends only on V_j , not on V_{medial} or V_{motor} . The shift of the positive limb between the curves for Δ h.p. = -30 mV (●) and Δ h.p. = +30 mV (◆) is not typical and we do not consider it significant.

Second, when extrapolated to time $t = 0$, I_j for the records at Δ h.p. = +30 mV (▲) produces an approximately linear instantaneous I/V plot, except for a slight bend near the holding current. This bend is a common result which may be due to imperfect extrapolation through fast kinetic components.

Third, there is a decreased resistance in the recovery I_j (■) pathway at negative V_j s. A progressive loss of rectification is a consistent feature of all experiments, usually terminating with a linear steady-state I/V . This may represent either the appearance of a parallel conductance pathway or the failure of a fraction of channels to close. The latter could argue for the gradual washing away of a soluble gating component by the perfusion system¹⁰, but other possible causes are hydrostatic pressure and chemical interactions with the perfusate. A similar loss of rectification has been reported sometimes to accompany transient pH uncoupling of intact preparations⁸. Whatever the

cause, the gating mechanism is extremely labile and the maximum synaptic resistance *in vivo* may actually be much higher than reported.

It is probable that the gating kinetics of this synapse have not been demonstrated previously because at room temperature gating is so fast that it appears instantaneous with respect to the capacitive charging of the membrane. Figure 4 shows I_j recorded from the medial axon of a voltage-clamped preparation in response to stepping the motor axon ± 85 mV while warming from 9.4°C to 25°C (Fig. 4a) and then cooling back to 10.3°C (Fig. 4b). The time constant of decay of I_j decreased dramatically as temperature increased, so that at 25°C it could no longer be resolved.

The temperature dependence of channel gating is obviously quite large. We have not yet determined an adequate function for fitting curves to these data. Approximating them by simple exponential functions, with no constraints on the initial jump, yields an effective temperature coefficient (Q_{10}) of roughly 11 for the time constant of decay over the temperature range studied. This is a surprisingly high figure in comparison to that of a typical gated sodium or potassium channel, which is thought to have a Q_{10} of about 3 (ref. 11). We are continuing investigations into possible explanations for this high Q_{10} . The change in positive current magnitude is consistent with a Q_{10} of 1.2, a typical value for a large aqueous pore¹².

Our data suggest that the conductance of the rectifying motor giant synapse of the crayfish is controlled by the voltage-dependent gating of gap-junction channels that open and close with a finite time constant, and that the instantaneous I/V plot is nearly linear before significant deactivation takes place. In comparison with the few other voltage-gated gap-junctions that have been described, this synapse resembles amphibian embryonic junctions⁹, and is unlike insect salivary gland junctions¹³ in that it is controlled by the transjunctional, not transmembrane, potential. However, it differs from the embryonic junctions in two significant ways: it is asymmetrical and it is much faster. Embryonic junctions shut off in either polarity⁹ and have time constants of several hundred milliseconds at room temperature¹⁴.

It has been proposed that the gating of the embryonic junctions is due to simple conformational changes in a part of the channel protein within the applied electric field, and that its symmetry arises from two such gates in series, one in each hemichannel¹⁴. Because of the high Q_{10} of the motor synapse, its gating mechanism is probably not so simple. In terms of this model, however, the crayfish junction would contain only one much faster gate, located in one hemichannel. This would probably be on the motor side, since the septal junctions coupling pairs of lateral axons are not voltage dependent¹⁵.

This work was supported by NIH grants NRSA NS06452 and HL31299. We thank Drs Lloyd Barr, Richard T. Mathias and Paul R. Adams for reading the manuscript and Ms Lucille Betti for preparing the figures.

Received 24 March; accepted 3 June 1986.

1. Furshpan, E. J. & Potter, D. D. *Nature* **180**, 342-343 (1957).
2. Furshpan, E. J. & Potter, D. D. *J. Physiol., Lond.* **145**, 289-325 (1959).
3. Bennett, M. V. L. in *Handbook of Physiology, The Nervous System* Vol. 1, 357-416 (Williams & Wilkins, Baltimore, 1977).
4. Cole, K. S. *Physiol. Rev.* **45**, 339-379 (1965).
5. Margiotta, J. F. & Walcott, B. *Nature* **305**, 52-55 (1983).
6. Giaume, C. & Korn, H. *Science* **220**, 84-87 (1983).
7. Giaume, C. & Korn, H. *J. Physiol., Lond.* **356**, 151-167 (1984).
8. Giaume, C. & Korn, H. in *Gap Junctions* (eds Bennett, M. V. L. & Spray, D. C.) 367-379 (Cold Spring Harbor Laboratory, New York, 1985).
9. Spray, D. C., Harris, A. L. & Bennett, M. V. L. *Science* **204**, 432-434 (1979).
10. Johnston, M. F. & Ramon, F. *J. Physiol., Lond.* **317**, 509-518 (1981).
11. Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **116**, 473-496 (1952).
12. Brink, P. R. in *Gap Junctions* (eds Bennett, M. V. L. & Spray, D. C.) 123-138 (Cold Spring Harbor Laboratory, New York, 1985).
13. Obaid, A. L., Socolar, S. J. & Rose, B. *J. Membrane Biol.* **73**, 69-89 (1983).
14. Harris, A. L., Spray, D. C. & Bennett, M. V. L. *J. gen. Physiol.* **77**, 95-117 (1981).
15. Johnston, M. F. & Ramon, F. *Biophys. J.* **39**, 115-117 (1982).
16. Glantz, R. M. & Viancour, T. *J. Neurophysiol.* **50**, 1122-1142 (1983).
17. Bezani, F. & Armstrong, C. M. *J. gen. Physiol.* **60**, 588-608 (1972).

Purified dihydropyridine-binding site from skeletal muscle t-tubules is a functional calcium channel

Veit Flockerzi*, Hans-Jochem Oeken*, Franz Hofmann*†, Dieter Pelzer‡, Adolfo Cavalié‡ & Wolfgang Trautwein‡

* Physiologische Chemie and ‡ II Physiologisches Institut, Medizinische Fakultät der Universität des Saarlandes, D-6650 Homburg-Saar, FRG

Many excitable cells contain at least two different voltage-dependent Ca channels (L- and T-type)¹⁻³. The cardiac, slow, L-type Ca channel is further modulated by cyclic AMP-dependent phosphorylation⁴⁻⁷, which increases the probability of it being open^{5,7}, and is readily blocked by Ca channel blockers including dihydropyridines and phenylalkylamines⁸⁻¹¹. The tritiated congeners of these blockers bind *in vitro* to sites which have the same pharmacological characteristics as those observed *in vivo*¹²⁻¹⁴, that is, stereospecific and allosteric interaction between distinct sites. The dihydropyridine-binding site purified from skeletal muscle t-tubules¹⁵ contains three peptides of relative molecular mass (M_r) 142,000 (142K), 56K and 31K. The cAMP kinase incorporates one mol phosphate per mol of the 142K peptide and binding of (+)PN-200/110, a potent Ca antagonist, is allosterically affected by D-*cis*-diltiazem and verapamil. The purified dihydropyridine-receptor complex has also been incorporated into phospholipid bilayer membranes. Here, we show for the first time that the complex can be reconstituted to form a functional 20-pS Ca channel that retains the principal regulatory, biochemical and pharmacological properties of membrane-bound L-type Ca channels.

We purified the nitrendipine receptor ~500-fold to an apparent density of 2 nmol per mg protein. Gel electrophoresis of the purified receptor in the absence of a reducing agent yielded three major peptides of apparent M_r 142K, 56K and 31K. Scans of the silver-stained gels indicate that 90% of the stain is localized with these peptides. Gel electrophoresis in the presence of a reducing agent yielded three additional peptides of M_r 122K, 26K and 22K (Fig. 1a). The significance of these additional bands is unclear. The 122K peptide could be part of the nitrendipine-binding site because azidopine, a photoaffinity label for this site, labels two peptides of similar M_r in the intact membrane¹⁷. The 142K and 56K peptides are phosphorylated by cAMP kinase (Fig. 1b), in agreement with previous reports^{18,19}. The phosphorylation of the 56K peptide in intact membranes has also been observed^{18,19}, but only one of these groups has reported the phosphorylation of the 142K peptide¹⁹. We tested the apparent specificity of the phosphorylation in the presence of a low concentration of cAMP kinase (7 nM) and ATP (3.1 μ M). The peptides had different initial phosphorylation rates under these conditions (Fig. 1b, inset). In the presence of 0.1 mM ATP, 1.02 ± 0.1 and 0.12 ± 0.01 mol phosphate were incorporated per mol 142K and 56K peptide, respectively. This suggests that the phosphorylation of the 142K peptide is functionally important. (The 122K and 31K peptides were not phosphorylated significantly.)

The receptor purified in the presence of tritiated nitrendipine contained variable amounts of this ligand. This probably prevented the determination of regular binding isotherms for nitrendipine, but a regular binding isotherm was obtained with tritiated (+)PN-200/110 (Fig. 2a), a dihydropyridine with a 10-fold higher affinity for the receptor than nitrendipine. Scatchard analysis yields a straight line with apparent $B_{\max}$ and K_D values

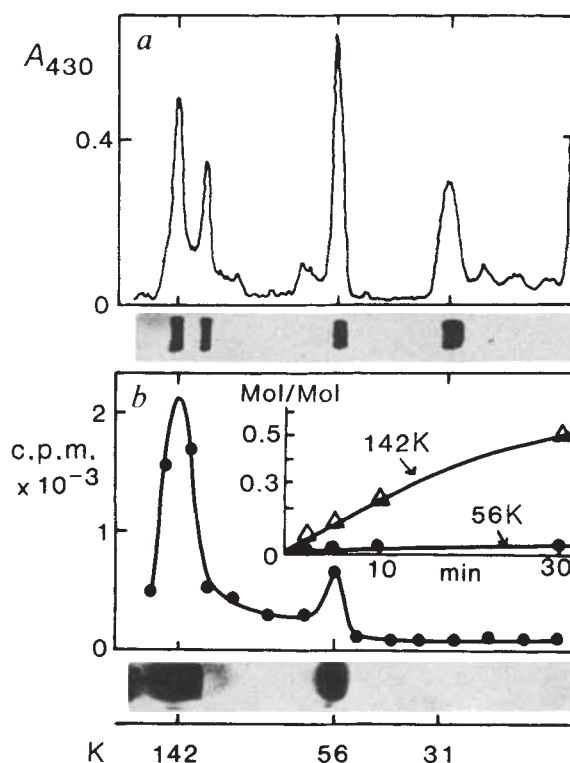


Fig. 1 Analysis of the purified nitrendipine-binding site on a 7.5% SDS gel. Purified receptor (150 ng) was boiled in the presence of 1% SDS and 40 mM dithiothreitol (DTT). a, Densitometric scan of a silver-stained gel. The areas under the 142, 122, 56 and 31K peptides have a ratio of 1.7:1.0:2.1:1.0. b, We phosphorylated 150 ng purified receptor by 7 nM pure catalytic subunit of cAMP kinase in the presence of 3.1 μ M 32 P-ATP (86 c.p.m. fmol^{-1}). The reaction was stopped after various times by boiling the sample in the presence of 1% SDS and 40 mM DTT. The gel was stained, dried and exposed to a Kodak X-Omatic film (lower part of b). The gel was then cut into 5-mm pieces which were dissolved in H_2O_2 and counted (58% efficiency) in a toluene-based scintillator. The relative ratio of stained areas was used for the molar calculation of the inset. The receptor was purified as described¹⁵ with the exception that the DEAE-Sephadex column was replaced by a TSK DEAE-5PW column. The gel pattern and binding characteristics of the purified receptor did not change if the receptor was kept at 4°C for 5 days. The 22 and 26K bands are only visible in the scan in a.

of 2.8 nmol mg^{-1} and 11.2 nM, respectively. A virtually identical K_D value of 9.8 nM and an ~10-fold lower value has been reported for the solubilized¹² and membrane-bound²⁰ receptor, respectively. The density of 2.8 nmol mg^{-1} protein suggests that the receptor is at least 60% pure because the M_r of the membrane-bound receptor has been reported to be between 180K and 210K^{12,21}.

PN-200/110 does not bind to the purified receptor in the presence of 1 mM EGTA but binds stereospecifically in the presence of calcium (Fig. 2b). The binding of (+)PN-200/110 is regulated allosterically by the benzothiazepine diltiazem and the phenylalkylamine verapamil (Fig. 2c). D-*cis*-diltiazem stimulates the binding of (+)PN-200/110, whereas L-*cis*-diltiazem added at the same concentration inhibits dihydropyridine binding. A partial inhibition of (+)PN-200/110 binding is obtained by verapamil. Verapamil also suppresses the stimulatory effect of D-*cis*-diltiazem (apparent EC_{50} = 10 μ M) with an apparently identical IC_{50} value of 10 μ M. These opposite effects of D-*cis*-diltiazem and verapamil have been reported previously with the membrane-bound^{12,14} and solubilized receptor^{22,23}, suggesting that the purified nitrendipine receptor contains sites specific for dihydropyridines, phenylalkylamines and probably diltiazem.

† To whom correspondence should be addressed.

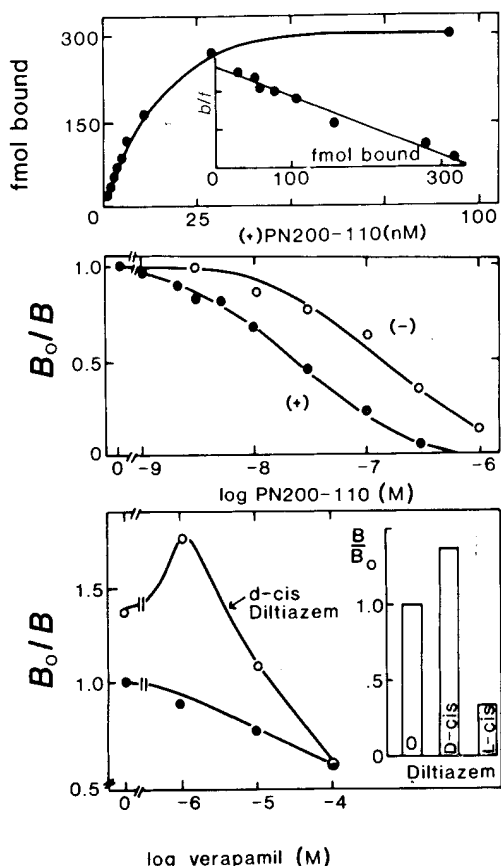


Fig. 2 Specific binding and allosteric regulation of (+)PN-200/110. *a*, Purified receptor was incubated with up to 2 nM ^3H -(+)-PN-200/110 and the indicated concentrations of unlabelled (+)PN-200/110. *b*, 1.7 nM ^3H -(+)-PN-200/110 was incubated with the indicated concentrations of (+) (●) and (-) (○) PN-200/110. *c*, ^3H -(+)-PN-200/110 (3.2 nM) was incubated in the absence and presence of 100 μM D- and L-cis diltiazem (see inset) and in the absence (●) and presence (○) of 100 μM D-cis diltiazem and various concentrations of (+)verapamil. The receptor concentration and incubation temperature were 125 ng and 4 °C (*a*, *b*) and 48 ng and 37 °C (*c*). Binding was measured as described elsewhere¹⁵. The amount bound is given in *a* as fmol per assay volume (40 μl). The receptor used in *c* was purified in the absence of D-cis diltiazem using (+)PN-200/110 as tritiated ligand. B_0 was 1,800 c.p.m. and nonspecific binding measured with 2 μM (+)PN-200/110 was 25% of the total binding.

Our results suggest that purification of the nitrendipine receptor does not alter its biochemical and pharmacological properties. However, we did not know whether the purified protein complex could be reconstituted to form a functional Ca channel. So far, reconstitution of Ca channel activity has been achieved only with crude solubilized t-tubular membrane fractions from rabbit and rat skeletal muscle^{24,25} and crude sarcolemmal vesicles from bovine and porcine cardiac muscle^{25,26}. We used the contact method¹⁶ to form solvent-free phospholipid bilayer membranes at the tip of glass patch pipettes. Figure 3A shows spontaneous single-channel openings of an incorporated purified binding site when a constant potential gradient is applied across the bilayer membrane. At an intra-pipette potential (V_{pip}) of -35 mV, the elementary current amplitude, I , is -0.75 pA (Fig. 3Aa). When V_{pip} is held at +35 mV, the channel carries the same amount of current in the opposite direction (Fig. 3Ab). The I - V_{pip} relation is linear for V_{pip} of ± 165 mV without distinct signs of rectification; the single-channel conductance is ~ 20 pS (Fig. 3Ac). Similar values have been observed in bilayer recordings of single Ca channels from solubilized skeletal muscle t-tubular and cardiac sarcolemmal

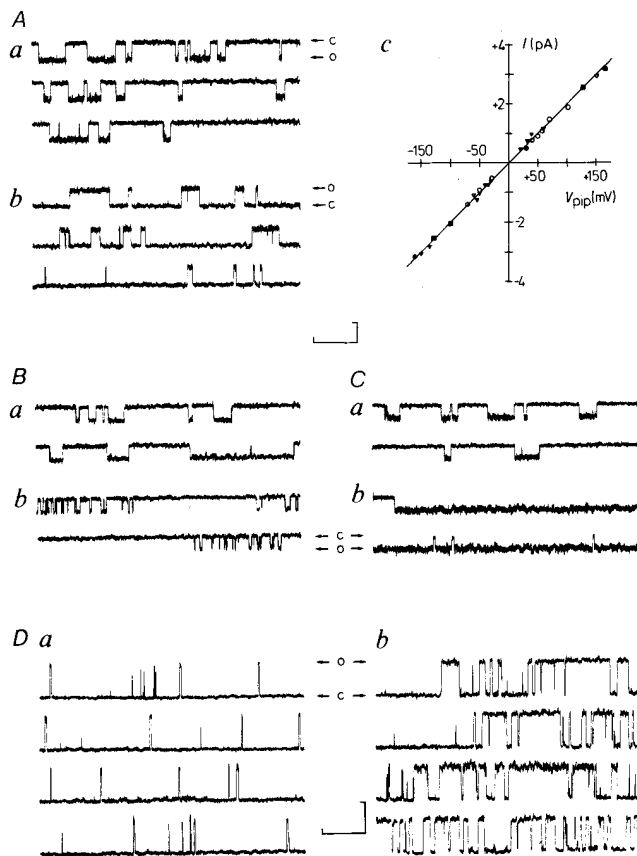


Fig. 3 *A*, Electrophysiology; *B*, *C*, pharmacology; *D*, biochemical regulation of the purified receptor. *A*, Steady-state recordings of spontaneous single-channel openings at a V_{pip} of -35 mV (*a*) and +35 mV (*b*), and the dependence of the elementary current amplitude (I) on V_{pip} (*c*). Each point in *c* represents the mean of >50 single-channel current amplitude measurements at their respective steady V_{pip} . The gradient indicates a single-channel conductance of 20 pS (in symmetrical 90 mM Ba). Scale bars, 1 pA (vertical), 0.5 s (horizontal). *B*, *C*, Control single-channel current records at a steady V_{pip} of -30 mV (*Ba*; *Ca*; control) and 15 min after bath application of 25 μM gallopamil (*Bb*) or 2.5 min after the addition of 10 μM BAY K 8644 (*Cb*). *D*, Single Ca channel currents at a steady V_{pip} of +165 mV before (*a*) and 1 min after (*b*) bath application of 1.25 μM C subunit of the cAMP kinase, 0.3 mM MgCl_2 and 1 mM ATP- γ -S. Scale bars, 3 pA (vertical), 0.5 s (horizontal). *A*-*D* show different experiments with distinct receptor preparations, but *a* (control) and *b* (experiment) of *B*-*D* were always recorded from the same bilayer. Elementary currents were digitized at 1 kHz and low-pass filtered at 300 Hz. Open-state probabilities (P_0) were obtained by time-averaging 15-min-long steady single-channel current records and dividing the time average by the single-channel I . Closed and open state of the channel is indicated by c and o, respectively.

Methods. We added phospholipids (PE:PS:cholesterol at a weight ratio of 2.8:1.2:1) to the purified receptor at a final concentration of 10 mg ml^{-1} and dialysed the sample against a 5 mM HEPES buffer, pH 7.4, containing 100 mM KCl, 50 mM NaCl and 1.5 mM CaCl_2 for 16 h at 4 °C. Dialysis was continued for a further 2 $\times$ 17 h against the same buffer in the absence of CaCl_2 . Dialysed liposomes were stored in small aliquots at -80 °C. Phospholipid bilayer membranes were formed from a mixture of bovine heart phosphatidylethanolamine (PE), bovine brain phosphatidylserine (PS) and cholesterol at a PE:PS:cholesterol ratio of 70:15:15 in *n*-hexane (1 mg lipid per ml). For monolayers, 5-8 μl lipid solution was spread at the surface of a 1.5-ml bath. The solution on both sides of the lipid bilayer contained 90 mM BaCl_2 and 5 mM HEPES (pH 7.3, 20-23 °C). The membrane potential across the bilayer was changed by clamping the V_{pip} with respect to the bath. According to the recording configuration and the convention of the amplifier used (L/M-EPC 7, List-Medical Electronic, Darmstadt), a positive current is defined as net cation flow from the pipette into the bath. Anionic current flow through incorporated proteins was excluded by the positive shift of the current reversal potential (+19 mV) on decreasing the BaCl_2 concentration in the pipette to 20 mM. Incorporation was achieved by bath addition of 10-40 μl of the liposome suspension corresponding to a total of 0.01-0.1 pmol binding sites. Data were analysed off-line. The currents were filtered with a 4-pole Bessel filter and digitized with a 12-bit A/D converter. The cut-off frequency (-3 dB) of the filter was usually adjusted to one-third of the sampling frequency.

membranes^{24,25}. Furthermore, this conductance is in agreement with results from cell-attached L-type Ca channel recordings²⁷.

The elementary currents through the 20-pS channel are blocked by the phenylalkylamine Ca channel blocker gallopamil. Bath application of 25 μ M gallopamil (final concentration) shortens the open-channel lifetimes ($V_{pip} = -30$ mV); consequently, the channel is usually closed (Fig. 3B). The gallopamil-induced flickering of channel activity in combination with the increased number of openings per unit time can be explained by rapid blocking-unblocking transitions, as described for the L-type Ca channel block by gallopamil in cardiac ventricular cells¹¹. At a V_{pip} of -30 mV, the open-state probability, P_o , of the channel is reduced from 0.29 ± 0.09 to 0.14 ± 0.03 by 14 μ M gallopamil ($n = 3$), and from 0.22 ± 0.05 to 0.02 ± 0.01 by 25 μ M gallopamil ($n = 4$). The ED_{50} of ~ 14 μ M corresponds closely to the ED_{50} of gallopamil for Ca channel block in frog skeletal muscle²⁸. The dihydropyridine Ca agonist BAY K 8644 (10 μ M final bath concentration) at the same V_{pip} (-30 mV) increases channel P_o from 0.2 ± 0.01 to 0.81 ± 0.08 ($n = 2$) by inducing long-term channel openings (Fig. 3C) as reported for L-type Ca channels in native cell membranes^{8,29,30}.

Phosphorylation of the receptor complex by cAMP kinase prolongs the open-channel lifetimes and shortens the shut intervals between channel openings and/or groups of openings (Fig. 3D). In two experiments, P_o of the channel increases from 0.05 ± 0.02 to 0.57 ± 0.11 . In three other experiments with higher control open-state probability, phosphorylation enhances the P_o value to a lesser extent from 0.25 ± 0.11 to 0.74 ± 0.07 . These findings also agree with the data on cAMP-dependent phosphorylation of the L-type Ca channel^{4-7,31,32}.

We did not obtain single-channel current fluctuations with similar electrophysiological, biochemical and pharmacological properties by adding heat-inactivated receptor; detergent-containing buffer treated as the receptor complex; cAMP kinase with or without Mg-ATP; Mg or ATP alone; or phospholipids alone. Channel-like elementary events with conductances larger or smaller than 20 pS, observed singly or in combination after addition of the liposome suspension to the bath, are insensitive to gallopamil up to a concentration of 2.5 mM and do not respond to cAMP-dependent protein phosphorylation. Whether or not these single-channel current fluctuations can be explained by incorporation of single peptides that lack the regulatory properties of the complete receptor complex, by aggregation of several receptor complexes and/or peptides forming nonspecific channels, by different channels with different pharmacological properties or by combinations of these remains to be seen.

Note that first, as the binding to purified channel proteins is still regulated allosterically, it seems that regulation is a property of the proteins and does not depend on the phospholipid environment. Second, we and others²⁵ observed that a decrease in Ca channel activity occurs in phospholipid bilayer membranes after incorporation, either as a gradual decrease of P_o or as a sudden irreversible channel closure. But the occurrence of spontaneous voltage-induced openings of Ca channels for a longer period after incorporation in the absence of phosphorylating agents (>2 h in our longest recordings) suggests either that a population of phosphorylated channels has been purified or that phosphorylation is not an absolute requirement for channel opening³². Third, phosphorylation of the 56K peptide may not be important for the cAMP-mediated modulation of Ca channel gating, which agrees with the observation that the purified nitrendipine receptor consists of only a 142K and a 31K peptide³³. Our results do not yet reveal which peptides are necessary for the reconstitution of a completely regulated Ca channel and whether all dihydropyridine-binding sites in intact skeletal muscle³⁴ are indeed functional Ca channels.

We thank Dr McDonald for correcting the final version of the manuscript, Mrs Nagel and Ms Bastuck for technical assistance, Mrs Siepmann for graphical work and Mrs Sommer for typing the manuscript. We also thank Drs Hollemann and

Kretzschmar (Knoll AG), Drs Lindemann and Friedli (Sandoz AG) and Drs Satzinger and Marmé (Gödecke AG) for the compounds used here. This work was supported by grants from Deutsche Forschungsgemeinschaft (SFB 246), Thyssen Stiftung and Fonds der chemischen Industrie.

Received 17 March; accepted 6 June 1986.

1. Carbone, E. & Lux, H. D. *Nature* **310**, 501-502 (1984).
2. Bean, B. P. *J. gen. Physiol.* **86**, 1-30 (1985).
3. Nilius, B., Hess, P., Lansman, J. B. & Tsien, R. W. *Nature* **316**, 443-446 (1986).
4. Osterrieder, W. *et al. Nature* **298**, 576-578 (1982).
5. Cachelin, A. B., de Peyer, J. E., Kokubun, S. & Reuter, H. *Nature* **304**, 462-464 (1983).
6. Bean, W. P., Nowicky, M. C. & Tsien, R. W. *Nature* **307**, 371-375 (1984).
7. Brum, G., Osterrieder, E. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **401**, 111-118 (1984).
8. Hess, P., Lansman, J. B. & Tsien, R. W. *Nature* **311**, 538-544 (1984).
9. Sanguinetti, M. C. & Kass, R. S. *Circulation Res.* **55**, 336-348 (1984).
10. McDonald, T. F., Pelzer, D. & Trautwein, W. *J. Physiol., Lond.* **352**, 217-241 (1984).
11. Cavalié, A., Pelzer, D. & Trautwein, W. *J. Physiol., Lond.* **358**, 59P (1985).
12. Glossmann, H., Ferry, D. R., Goll, A., Striessnig, J. & Zernig, G. *Arzneim-Forschung/Drug Res.* **35**, 1917-1935 (1985).
13. Snyder, S. H. & Reynolds, I. J. *New Engl. J. Med.* **313**, 995-1002 (1985).
14. Ruth, P., Flockerzi, V., v. Nettelbladt, E., Oeken, J. & Hofmann, F. *Eur. J. Biochem.* **150**, 313-322 (1985).
15. Curtis, B. M. & Catterall, W. A. *Biochemistry* **23**, 2113-2118 (1984).
16. Hanke, W., Methfessel, C., Wilmsen, U. & Boheim, G. C. *Bioelectrochem. Bioenerget.* **12**, 329-339 (1984).
17. Ferry, D. R., Kämpf, K., Goll, A. & Glossmann, H. *EMBO J.* **4**, 1933-1940 (1985).
18. Curtis, B. M., Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2528-2532 (1985).
19. Hosey, M. M., Borsotto, M. & Lazdunski, M. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
20. Borsotto, M., Barhanin, J., Norman, R. I. & Lazdunski, M. *Biochem. biophys. Res. Commun.* **122**, 1357-1366 (1984).
21. Norman, R. I., Borsotto, M., Fossett, M., Lazdunski, M. & Ellory, J. C. *Biochem. biophys. Res. Commun.* **111**, 878-883 (1983).
22. Borsotto, M., Norman, R. I., Fossett, M. & Lazdunski, M. *Eur. J. Biochem.* **142**, 449-455 (1984).
23. Ruth, P., Flockerzi, V., Oeken, H.-J. & Hofmann, F. *Eur. J. Biochem.* **155**, 613-620 (1986).
24. Affolter, H. & Coronado, R. *Biophys. J.* **48**, 341-347 (1985).
25. Rosenberg, R. L., Hess, P., Tsien, R. W., Smilowitz, H. & Reeves, J. P. *Science* **213**, 1564-1566 (1986).
26. Ehrlich, B. E., Schen, C. R., Garcia, M. L. & Kaczorowski, G. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 193-197 (1986).
27. Cavalié, A., Ochi, R., Pelzer, D. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **398**, 284-297 (1983).
28. Palade, P. T. & Almers, W. *Pflügers Arch. ges. Physiol.* **405**, 91-101 (1985).
29. Kokubun, S. & Reuter, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4824-4827 (1984).
30. Ochi, R., Hino, N. & Niimi, Y. *Proc. Japan Acad.* **60**, B, 153-156 (1984).
31. Kameyama, M., Hofmann, F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **405**, 285-293 (1985).
32. Kameyama, M., Hescheler, J., Hofmann, F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* (in the press).
33. Borsotto, M., Barhanin, J., Fossett, M. & Lazdunski, M. *J. biol. Chem.* **260**, 14255-14263 (1985).
34. Schwartz, L. M., McClesley, E. W. & Almers, W. *Nature* **314**, 747-751 (1985).

Activation of two signal-transduction systems in hepatocytes by glucagon

Michael J. O. Wakelam, Gregory J. Murphy, Victor J. Hruby* & Miles D. Houslay

Molecular Pharmacology Group, Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, UK

The ability of glucagon to stimulate glycogen breakdown in liver played a key part in the classic identification of cyclic AMP and hormonally stimulated adenylate cyclase¹. But several observations indicate that glucagon can exert effects independent of elevating intracellular cAMP concentrations²⁻⁷. These effects are probably mediated by an elevation^{8,9} of the intracellular concentration of free Ca^{2+} although the mechanism by which this occurs is unknown. We show here that glucagon, at the low concentrations found physiologically, causes both a breakdown of inositol phospholipids and the production of inositol phosphates. Indeed, we show that the glucagon analogue, (1-N- α -trinitrophenylhistidine,12-homo-arginine)glucagon (TH-glucagon), which does not activate adenylate cyclase or cause any increase in cAMP in hepatocytes yet can fully stimulate glycogenolysis, gluconeogenesis and urea syn-

* Department of Chemistry, University of Arizona, Tucson, Arizona 85721, USA

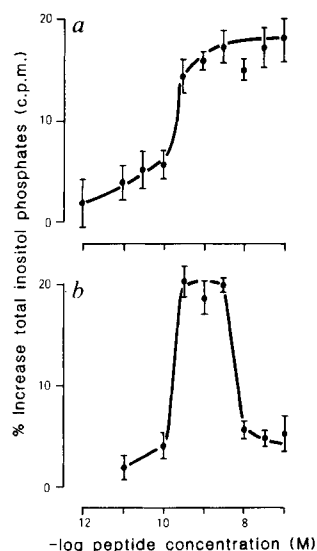


Fig. 1 Dose-response curves of the effect of glucagon and TH-glucagon on inositol phosphate production. *a*, Response to TH-glucagon; *b*, response to glucagon. Results are expressed as means $\pm$ s.e.m. of the percentage increase ($\delta\%$) in radioactivity in inositol phosphates above that of the control value (no hormone). Data were pooled from experiments using seven different cell preparations ($n=21$). Typically the control values were for lipids 8,500–10,000 c.p.m. and for the inositol phosphates 200–300 c.p.m. Over the concentration range 5×10^{-11} – 10^{-7} M, TH-glucagon stimulated inositol phosphate production significantly ($P < 0.001$) and at 10^{-11} M ($P \leq 0.05$). Glucagon achieved a significant stimulation of inositol phosphate production over the concentration range 10^{-10} M to 5×10^{-8} M ($P < 0.001$).

Methods. Rat hepatocytes were prepared as described in Fig. 3 legend. They were incubated in Krebs–Henseleit buffer containing 1% (w/v) bovine serum albumin, 10 mM glucose and $5 \mu\text{Ci ml}^{-1}$ [^3H]myo-inositol for 90 min with constant gassing using a 95% O_2 :5% CO_2 mixture. The cells were then washed and resuspended in the same buffer in the absence of inositol, and incubated for 10 min before the addition of LiCl to a final concentration of 10 mM, whereupon incubation was continued for 15 min. The cells were then dispensed into plastic vials containing glucagon or TH-glucagon at the stated final concentration, gassed (O_2/CO_2), capped and incubated for 30 min. All incubations were performed in a 37°C shaking water bath. Incubations were terminated by the addition of CHCl_3 :methanol (1:2). Following the addition of CHCl_3 and water to generate two phases, the radioactivity in the total inositol phosphates was determined by batch chromatography on Dowex 1×8 formate resins²³. Radioactivity in the lipids was determined on a portion of the lower CHCl_3 phase. Synthesis and purification of TH-glucagon was performed as described previously²⁴.

thesis¹⁰, stimulates the production of inositol phosphates. This stimulation of inositol phospholipid metabolism by low concentrations of glucagon provides a mechanism^{11,12} whereby glucagon can exert cAMP-independent actions on target cells. We suggest that hepatocytes possess two distinct receptors for glucagon, a GR-1 receptor coupled to stimulate inositol phospholipid breakdown and a GR-2 receptor coupled to stimulate adenylate cyclase activity.

Figure 1 shows that both glucagon and TH-glucagon cause a dose-dependent increase in the production of inositol phosphates in intact hepatocytes with K_a values of 0.25 ± 0.02 nM and 0.3 ± 0.01 nM, respectively. The dose-dependence of this effect is very similar to that seen for the glucagon-mediated desensitization⁷ of adenylate cyclase ($K_a = 0.4$ nM) and the ability of glucagon to elevate^{8,9} intracellular free Ca^{2+} ($K_a = 0.2$ – 0.3 nM). The dose-response curve for the glucagon-stimulated production of inositol phosphates is clearly biphasic, with inhibition occurring at higher glucagon concentrations. The K_a value for this inhibitory phase was found to be 8.5 ± 0.2 nM,

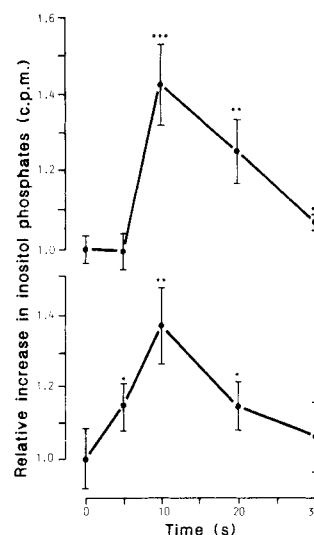


Fig. 2 Effect of glucagon and TH-glucagon on InsP_3 (*a*) and InsP_2 (*b*) levels in hepatocytes. Results from one experiment, which is typical of two which gave similar results, are expressed as means $\pm$ s.d. of the relative increase in radioactivity in the inositol phosphates compared with that seen at time = zero ($t=0$). In each case $n=3$. The radioactivity at $t=0$ was 954 ± 80 c.p.m. for InsP_3 and $4,986 \pm 170$ for InsP_2 . The lipids, extracted using acidified CHCl_3 :MeOH from the perchloric acid pellet, contained $57,626 \pm 975$ c.p.m. Using a *t*-test to assess the significance of the elevated InsP_3 and InsP_2 concentrations then for data points marked as * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Those unmarked were not significantly different ($P > 0.05$) from the concentration at $t=0$.

Methods. Hepatocytes were prepared and the lipids labelled with [^3H]myo-inositol as described in Fig. 1 legend except that the final concentration was $50 \mu\text{Ci ml}^{-1}$. Hepatocytes were incubated, in the presence of 10 mM LiCl, with either glucagon or TH-glucagon (both 10^{-9} M) for the time stated. Incubations were terminated by the addition of 0.4 volumes of 10% (v/v) ice-cold perchloric acid. Following vortexing and centrifugation the supernatants were neutralized and applied to 1-ml Dowex 1×8 Formate columns. The inositol phosphates were separated²⁵ and radioactivity determined on portions of the eluates.

which is remarkably similar to the K_a value of 6.3 nM (Fig. 3) which we found for the ability of glucagon to stimulate adenylate cyclase. As TH-glucagon gives no evidence of such an inhibitory phase then it appears that the glucagon-stimulated inositol lipid response can be inhibited either as a result of the activation of the stimulatory guanine nucleotide regulatory protein N_s , by glucagon, or by elevation of intracellular cAMP concentrations. Others¹³ have demonstrated that elevation of cAMP concentrations does not inhibit the inositol lipid response to vasopressin in hepatocytes, implying that the former possibility is the most likely. Indeed, we found that the addition of dibutyryl cAMP to intact hepatocytes, at concentrations (10–100 μM) able to elicit biological effects⁶, failed to inhibit the ability of TH-glucagon (10^{-7} M) to increase total inositol phosphate production. However, treatment of hepatocytes with cholera toxin ($10 \mu\text{g ml}^{-1}$), under conditions which we have shown to activate maximally the functioning of N_s (30–60 min at 37°C)^{14,15}, causes a marked reduction ($\sim 85\%$) of the inositol phosphate response elicited by TH-glucagon (10^{-7} M). In contrast, incubation of hepatocytes with pertussis toxin under conditions (100 ng ml^{-1} for 1.5 h) which obliterate the functioning of the inhibitory guanine nucleotide regulatory protein N_i (ref. 16), augments ($\sim 25\%$ increase) the magnitude of the stimulation of the production of inositol phosphates by TH-glucagon.

Examination of the individual inositol phosphates produced by stimulation of hepatocytes with either glucagon or TH-glucagon shows (Fig. 2) that inositol trisphosphate (InsP_3) is

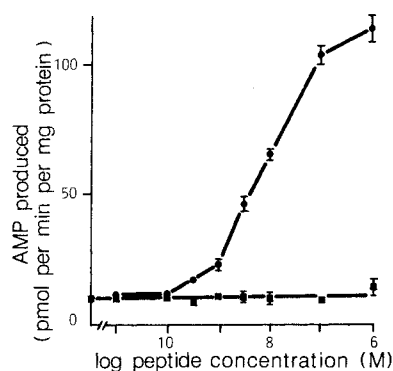


Fig. 3 TH-glucagon does not activate adenylate cyclase. Increasing concentrations of glucagon (●) activated adenylate cyclase with $K_a = 6.3 \pm 2.5$ nM. Only at concentrations $\geq 10^{-9}$ M does it increase cAMP concentrations significantly. Increasing concentrations of TH-glucagon (■) fails to produce any stimulation of the adenylate cyclase activity of isolated hepatocyte plasma membranes at any of the concentrations used. Error bars $\pm$ s.d. for three experiments ($n=3$) using triplicate determinations of activity. Basal adenylate cyclase activity was 10 pmol min per mg per protein.

Methods. Fed, male 200–250 g Sprague–Dawley rats were used in this and the other experiments described here. Hepatocytes were prepared as described previously⁷ using a collagenase digestion method which allowed for >90% viability as indicated by ATP content, glucagon-stimulated gluconeogenesis and trypan-blue exclusion. Cells ($3\text{--}5$ mg dry wt ml^{-1}) were preincubated at 37°C for 20–30 min before their use in experiments and medium and cells regularly gassed every 15 min with O_2/CO_2 (95/5%) mixture. A washed plasma membrane fraction was obtained as described previously^{7,14} and membranes, which were kept on ice, were used within 2 h of preparation. Adenylate cyclase was assayed as before²⁶ in a mixture of final concentrations of 1.5 mM ATP, 5 mM MgSO_4 , 10 mM theophylline, 1 mM EDTA, 7.4 mg phosphocreatine per ml, 1 mg creatine kinase per ml and 50 mM triethanolamine HCl adjusted to a final pH of 7.4 with KOH. The cAMP produced was measured as before⁷ with a binding assay using the purified regulatory subunit of bovine heart muscle cAMP-dependent protein kinase. The intracellular concentration of cAMP that occurred after a 5-min challenge of the hepatocytes with either glucagon (10^{-8} M) or TH-glucagon (10^{-8} M) was determined in the presence of isobutyl methylxanthine (1 mM) as described previously⁶. Basal concentration is 12.62 ± 1.60 pmol mg^{-1} dry wt cells; after glucagon challenge it is 310.98 ± 28.4 pmol mg^{-1} ; and after TH-glucagon challenge it is 13.70 ± 1.53 pmol mg^{-1} dry wt cells. Results are means $\pm$ s.e.m. from three cell preparations ($n=6$).

the first species to be generated. This response occurs within ~ 5 s of the addition of either glucagon or TH-glucagon. Elevated inositol biphosphate (InsP_2) concentrations are only observed 10 s after exposure to these hormones and no change in inositol monophosphate (InsP) is observed for the first 30 s. These effects of glucagon and TH-glucagon on inositol phospholipid metabolism are as expected from a 'classic' receptor-stimulated inositol phospholipid response^{11,12}. However, note that this effect is considerably smaller than the stimulation of inositol phospholipid metabolism elicited by vasopressin in hepatocytes^{13,17}. It is already known¹² that α_1 adrenoceptors in liver seem to produce a disproportionately large physiological response in relation to their rather small stimulation of inositol phospholipid metabolism. Thus our observations suggest a molecular mechanism for the reported actions of glucagon in eliciting small increases^{8,9} in both intracellular Ca^{2+} and in the production¹⁸ of diacylglycerol. We will show elsewhere that treatment of intact hepatocytes with TH-glucagon, vasopressin and angiotensin II can all elicit desensitization of glucagon-stimulated adenylate cyclase, a process mimicked¹⁹ by phorbol esters, presumably through protein C kinase activation and modification of the guanine nucleotide regulatory protein system in liver^{7,16}.

TH-glucagon exerts little effect on intact hepatocyte intracellular cAMP concentrations¹⁰, although in these experiments no inhibitor of cAMP phosphodiesterase activity was present. Because glucagon can activate hepatocyte cAMP phosphodiesterase activity⁶, it was unclear whether TH-glucagon failed to cause a net increase in intracellular cAMP concentrations by potentially stimulating such a phosphodiesterase rather than by failing to activate adenylate cyclase. Using both isolated hepatocyte plasma membranes (Fig. 3) and intact hepatocytes (see Fig. 3 legend), each assayed in the presence of a nonspecific cAMP phosphodiesterase inhibitor (theophylline or isobutyl methylxanthine), which obliterate >98% of phosphodiesterase activity⁶, we demonstrate clearly that TH-glucagon cannot stimulate adenylate cyclase activity. We also observe that TH-glucagon can inhibit the action of glucagon in stimulating adenylate cyclase in a simple competitive fashion (data not shown).

That TH-glucagon can elicit an inositol phospholipid response but not stimulate adenylate cyclase suggests that glucagon exerts its action on liver via two receptors rather than via a single receptor coupled to two signalling systems. The K_a value (0.25 nM) for the action of glucagon in stimulating the inositol phospholipid response is considerably lower than that observed for the ability of glucagon to activate adenylate cyclase in either isolated hepatocyte membranes ($K_a = 6.3 \pm 2.5$ nM; Fig. 3) or intact cells ($K_a = 1.2 \pm 0.1$ nM; ref. 12). Further support for the existence of two glucagon receptors is provided by binding studies using glucagon itself²⁰ and also with a synthetic glucagon analogue²¹, both of which identify two populations of glucagon receptors in the liver. This situation, with one receptor population coupled to stimulate inositol phosphate metabolism and the other to the stimulation of adenylate cyclase, is then analogous to several other multiple receptor systems such as adrenaline, functioning through α_1 - and β -adrenoceptors, and vasopressin, acting through V_1 and V_2 receptors^{12,13}. We suggest that rat hepatocytes have two functionally distinct glucagon receptors that show slightly different affinities for glucagon. By analogy with the vasopressin system in liver¹³ we suggest that the population of glucagon receptors coupled to inositol lipid breakdown are termed GR-1 receptors and those receptors that stimulate adenylate cyclase GR-2 receptors.

At physiological glucagon concentrations ($\sim 10^{-10}$ M) it is probable that a significant proportion of glucagon activity on cellular metabolic processes will be mediated via GR-1 receptors. But, at elevated glucagon concentrations, which are those used by most experimenters, the functioning of GR-2 receptors coupled to adenylate cyclase activation predominates. It is intriguing that these two receptors trigger events that attenuate each other's cellular function. GR-1 receptor occupancy leads to desensitization⁷ of glucagon-stimulated adenylate cyclase activity, as well as the production of a selective insulin-resistant state^{6,22}, and GR-2 receptor occupancy leads to N_g -mediated inhibition of glucagon-stimulated inositol phosphate production.

This study was supported by grants from the MRC, Scottish Home & Health Department, the California Metabolic Research Foundation and a Wellcome Trust Research Travel Grant to M.D.H. as well as USPHS grant AM21085 to V.J.H. We thank Miss Sandra Gardner for technical assistance and Dr Dev Trivedi for the preparation and purification of TH-glucagon.

Received 1 April; accepted 6 June 1986.

1. Sutherland, E. W. & Rall, T. W. *J. biol. Chem.* **232**, 1077–1091 (1958).
2. Birnbaum, M. J. & Fain, J. N. *J. biol. Chem.* **252**, 528–535 (1977).
3. Okajima, F. & Ui, M. *Archs Biochem. Biophys.* **175**, 549–557 (1976).
4. Cardenas-Tanus, R. & Garcia-Sainz, J. A. *FEBS Lett.* **143**, 1–4 (1982).
5. Khan, B. A., Bregman, M. D., Nugent, C. A., Hruby, V. J. & Brendel, K. *Biochem. biophys. Res. Commun.* **93**, 729–736 (1980).
6. Heyworth, C. M., Wallace, A. V. & Houslay, M. D. *Biochem. J.* **214**, 99–110 (1983).
7. Heyworth, C. M. & Houslay, M. D. *Biochem. J.* **214**, 93–98 (1983).
8. Sistaire, F. D., Picking, R. A. & Haynes, R. C. *J. biol. Chem.* **260**, 12744–12747 (1985).
9. Mauger, J.-P. & Claret, M. *FEBS Lett.* **195**, 106–110 (1986).
10. Corvera, S. et al. *Biochim. biophys. Acta* **804**, 434–441 (1984).

11. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
12. Downes, C. P. & Michell, R. H. *Molec. Aspects Cell Regul.* **4**, 2-56 (1985).
13. Creba, J. A., Downes, C. P., Hawkins, P. T., Brewster, G., Michell, R. H. & Kirk, C. J. *Biochem. J.* **212**, 733-747 (1983).
14. Houslay, M. D. & Elliott, K. R. F. *FEBS Lett.* **359**-363 (1979).
15. Heyworth, C. M., Whetton, A. D., Wong, S., Martin, R. B. & Houslay, M. D. *Biochem. J.* **228**, 593-603 (1985).
16. Heyworth, C. M., Hanski, E. & Houslay, M. D. *Biochem. J.* **222**, 189-194 (1984).
17. Thomas, A. P., Alexander, J. & Williamson, J. F. *J. biol. Chem.* **259**, 5574-5584 (1984).
18. Bockino, S. B., Blackmore, P. F. & Exton, J. H. *J. biol. Chem.* **260**, 14201-14207 (1985).
19. Heyworth, C. M., Wilson, S. R., Gawler, D. & Houslay, M. D. *FEBS Lett.* **187**, 196-200 (1985).
20. Sonne, O., Berg, T. & Christofferson, T. *J. biol. Chem.* **253**, 3203-3210 (1978).
21. Musso, G. F., Assoian, R. K., Kaiser, E. T., Kezdy, F. J. & Tager, H. S. *Biochem. biophys. Res. Commun.* **119**, 713-719 (1984).
22. Heyworth, C. M. & Houslay, M. D. *Biochem. J.* **214**, 547-552 (1983).
23. Berridge, M. J., Downes, C. P. & Hanley, M. R. *Biochem. J.* **206**, 587-595 (1982).
24. Bregman, M. D., Trivedi, D. & Hruby, V. J. *J. biol. Chem.* **255**, 11725-11733 (1980).
25. Berridge, M. J., Dawson, R. M. C., Downes, C. P., Heslop, J. P. & Irvine, R. F. *Biochem. J.* **212**, 473-482 (1983).
26. Houslay, M. D., Metcalfe, J. C., Warren, G. B., Hesketh, T. R. & Smith, G. A. *Biochim. biophys. Acta* **436**, 489-494 (1976).

***In vitro* synthesized bacterial outer membrane protein is integrated into bacterial inner membranes but translocated across microsomal membranes**

Makoto Watanabe, John F. Hunt* & Günter Blobel

Laboratory of Cell Biology, Rockefeller University, New York, New York 10021, USA

The LamB protein is an integral membrane protein of the outer membrane of *Escherichia coli*. We have now found that, when synthesized in an *E. coli* cell-free translation system supplemented with inverted vesicles derived from the *E. coli* inner membrane, LamB protein is integrated into the vesicle membrane as assayed by its resistance to extraction at alkaline pH. These data suggest that the inner membrane is the primary site for integration of LamB protein prior to subsequent sorting to the outer membrane. When synthesized in a wheat germ cell-free translation system supplemented with canine microsomal membranes, LamB protein is glycosylated at one or two cryptic sites, and surprisingly, it is translocated across instead of being integrated into the vesicle membrane. We suggest that the translocation machinery of the microsomal membrane, although able to recognize the signal sequence(s) of LamB, is unable to recognize its stop-transfer sequence(s), thereby yielding translocation instead of integration.

The disposition of LamB protein in the outer membrane of *E. coli* is 'polytopic'¹, that is, segments of the polypeptide are alternately exposed on both sides of the membrane. The protein is structurally and functionally related to the porins, the major protein constituents of the outer membrane of Gram-negative bacteria². These proteins are structurally distinct in that they contain β -sheet structure traversing the phospholipid bilayer and little or no α -helical structure²⁻⁴. Furthermore, these proteins are organized as trimers in the outer membrane^{2,5,6}, and trimerization is believed to be required for their activity⁷. The porins, including the LamB protein, act as relatively nonspecific channels, allowing free diffusion of small hydrophilic and ionic solutes across the Gram-negative outer membrane⁸⁻¹⁰. In addition, LamB protein functions as a carrier for maltose and maltodextrins containing up to eight glucose residues, catalysing facilitated diffusion of these species into the periplasmic space¹⁰⁻¹². Thus, LamB protein domains exposed to the periplasm interact with the maltose-binding protein¹²⁻¹⁴, while domains exposed to the extracellular milieu are involved in binding maltodextrins¹¹ as well as serving as the receptor for

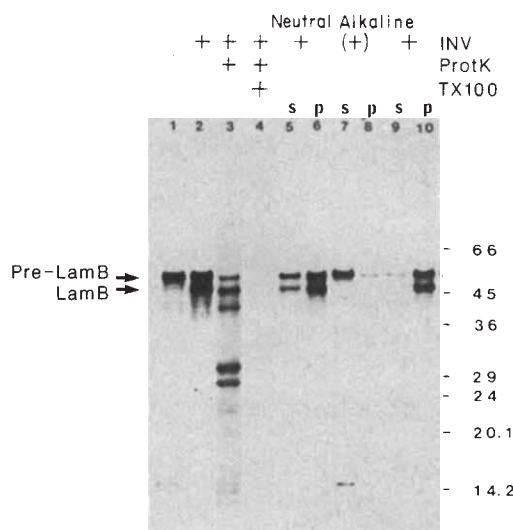


Fig. 1 LamB is integrated into the plasma membrane of *E. coli*. LamB mRNA was translated in a cell-free translation system of *E. coli* depleted of membranes. Where indicated, inverted vesicles (INVs) were present during translation. In lanes 7 and 8 INVs were added post-translationally together with NaOH. Post-translational incubation with proteinase K was carried out in the absence or presence of Triton X-100 (TX100). Post-translational sedimentation was carried out at neutral or alkaline pH and yielded a supernatant (s) and pellet (p) fraction. Numbers on the right indicate M_r of protein standards ($\times 10^3$). Arrows on the left indicate position of pre-LamB and mature LamB.

Methods. The SP64 plasmid containing the gene for LamB was a gift from Dr S. Benson. After linearization with *EcoRI*, transcription was performed as described previously¹⁹, and transcripts were purified by phenol/chloroform extraction followed by ethanol and LiCl precipitation. Translation was performed in a membrane-depleted high-speed supernatant (S-135) of an *E. coli* homogenate¹⁷; 12 μ Ci of [³⁵S]methionine and 200 ng of LamB mRNA were added to each 25- μ l reaction. Where indicated, inverted membrane vesicles (INVs)²⁰ were present at a final concentration of 1.0 $A_{280} \text{ ml}^{-1}$. Translations were carried out for 30 min at 37°C. Protease digestion was carried out in the presence of 5 mM CaCl_2 and 300 $\mu\text{g ml}^{-1}$ of proteinase K (ProtK) for 2 h at 4°C. Where indicated, Triton X-100 was present at a final concentration of 1%. For sedimentation at neutral pH, 25 μ l of the chilled reaction mixture was layered over a 50- μ l cushion of 0.5 M sucrose, 40 mM triethanolamine acetate (pH 7.5), 140 mM KOAc, 11 mM $\text{Mg}(\text{OAc})_2$, 20 mM NH_4OAc , 0.1 mM EDTA, 1 mM dithiothreitol; centrifugation was performed in an A100/18 rotor in the Beckman airfuge for 15 min at 30 p.s.i. (100,000g) at 4°C. After centrifugation, the entire supernatant (including the sucrose cushion) was removed and precipitated with an equal volume of ice-cold 20% trichloroacetic acid (TCA). The pellet fraction and TCA-precipitated samples were resuspended in electrophoresis sample buffer²⁹. Sedimentation at alkaline pH was carried out as previously described²³. Briefly, 25 μ l of the translation mixture was adjusted to pH 11-12 with 1.0 M NaOH and was layered over a sucrose buffer cushion adjusted to pH 11-12 with NaOH. Centrifugation was performed in an A100/18 rotor in the Beckman airfuge for 30 min at 30 p.s.i. at 4°C. Samples were analysed by SDS-polyacrylamide gel electrophoresis (PAGE) on slab gels containing a 10-15% acrylamide gradient and 0.5 SDS. The dried SDS gel was subjected to fluorography.

the bacteriophage λ (ref. 15). Like other outer membrane proteins, LamB protein is synthesized as a larger precursor with an amino-terminal signal sequence (25 residues) that is cleaved upon conversion to the mature form (421 residues)¹⁶.

We translated SP6 messenger RNA for the LamB protein, according to previously published methods¹⁷⁻¹⁹. In brief, we employed a cell-free translation system of *E. coli* that was either depleted or supplemented with inverted vesicles (INVs)²⁰ of the

* Present address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, USA.

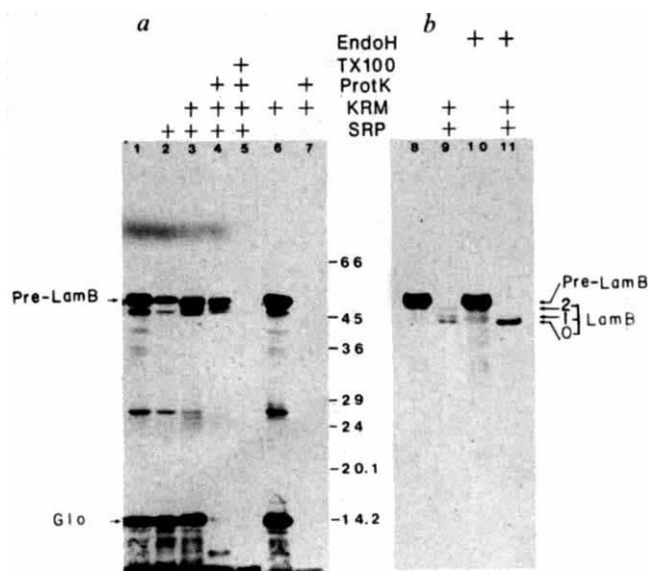


Fig. 2 LamB is protected from proteinase K digestion and glycosylated when segregated from a wheat germ cell-free translation by canine microsomal membranes. **a**, **b**, LamB mRNA (and in **a**, rabbit globin mRNAs (Glo) also) was translated in the wheat germ cell-free translation system. Where indicated, SRP and/or KRM were present during translation. Post-translational incubations were performed either with proteinase K (ProtK), in the absence or presence of Triton X-100 (TX100), or with Endo-H. Arrows indicate position of pre-LamB and LamB putatively containing 2, 1 or 0 Asn-linked, Endo-H-sensitive oligosaccharides.

Methods. 400 ng of LamB mRNA (and 1 µg of rabbit reticulocyte RNA, **a** only) were present in a 25-µl wheat germ translation mixture²⁹. Signal recognition particle (SRP) (80 U in **a**, 160 U in **b**) and 2 µl of salt-extracted microsomal membranes (KRM), both prepared from canine pancreas^{32,33}, were present in the lanes indicated. Translations were incubated for 60 min at 25 °C. Digestion with proteinase K (see Fig. 1) was performed for 30 min at 4 °C. Where indicated, Triton X-100 was present at a final concentration of 1%. Digestion with endo-*N*-acetylglucosaminidase H (Endo-H) was carried out as previously described³⁵. The dried SDS gel (see Fig. 1) was subjected to autoradiography.

E. coli inner membrane (Fig. 1). In the absence of INVs, only the precursor form of LamB was synthesized (Fig. 1, lane 1) whereas in the presence of INVs, there was synthesis also of the mature form (lane 2). The previously observed¹⁷ protection of the mature form (and of a fraction of the precursor form) suggested that LamB protein was either translocated into the lumen of the INVs or was integrated into the INV membrane in such a way that it was resistant to added proteinase K. To distinguish between these two possibilities we used two approaches. First, incubation with proteinase K was carried out for longer times than used previously. In these conditions a substantial amount of pre- and mature LamB was in fact converted to distinct smaller fragments (Fig. 1, lane 3). Protection of these fragments was due to their association with the membrane as demonstrated by their complete digestion in the presence of Triton X-100 (Fig. 1, lane 4). Thus, some domains of the LamB protein were accessible on the INV surface and other domains appeared to be integrated into the INV membrane and therefore inaccessible to protease.

In order to confirm this conclusion using an independent experimental approach, we carried out sedimentation of the LamB protein translation products at neutral and alkaline pH. As treatment at alkaline pH will break open membrane vesicles, translocated proteins will be extracted whereas membrane-

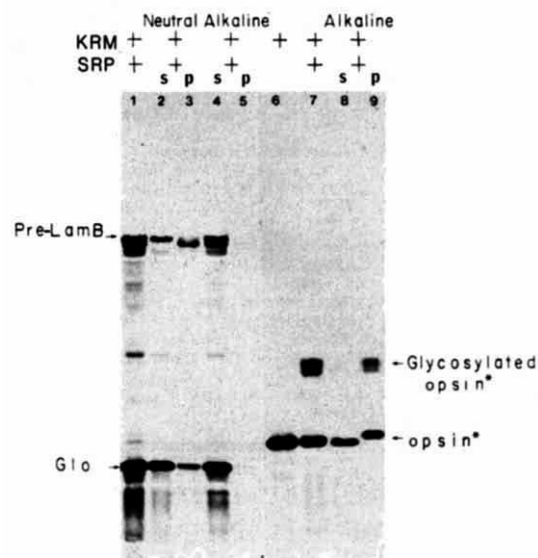


Fig. 3 LamB segregated by canine microsomal membranes is extractable at alkaline pH. mRNAs for LamB and rabbit globin mRNA (lanes 1-5) or mRNA for a derivative of bovine opsin³⁶ (opsin*) (lanes 6-9) were translated in the wheat germ cell-free system. Where indicated, SRP (80 U) and/or KRM (2 µl) were present during translation; post-translational sedimentation was carried out at neutral or alkaline pH and yielded a supernatant (s) or pellet (p) fraction. Arrows indicate position of pre-LamB, globin, opsin* and glycosylated opsin*.

Methods. As described in the legends to Figs 1 and 2. An SP65 plasmid containing a cDNA derivative of bovine opsin (pSF8)³⁶, referred to here as opsin*, was linearized with *HindIII*. 200 ng of opsin* mRNA was present in a 25-µl wheat germ translation. Sedimentation conditions at neutral and alkaline pH were described previously²³. The dried SDS gel was subjected to autoradiography.

integrated proteins will remain associated and sediment with the INV membrane²¹⁻²³. Indeed, the bulk of both pre- and mature LamB protein co-sedimented with the membranes after incubation at neutral (Fig. 1, lanes 5, 6) or alkaline pH (lanes 9, 10). The small amount of both forms found in the supernatant (Fig. 1, lane 5) was most probably due to incompletely sedimented membranes, as doubling the centrifugation time (15 min in lanes 5 and 6 against 30 min in lanes 9 and 10) led to an increased yield of both forms in the pellet fraction. However, not all pre-LamB that sedimented after alkali treatment was integrated. A control in which INVs were not present during translation but were added together with NaOH only after translation revealed co-sedimentation of a small amount of pre-LamB (Fig. 1, lanes 7, 8), suggesting that minor aggregation of pre-LamB might occur. Another control showed that maltose-binding protein, a periplasmic protein, was protected from protease digestion, sedimented at neutral pH, but was extracted at alkaline pH (data not shown). This means that it was translocated into the lumen of the INVs. Thus, from the results of both the protease digestion and alkaline extraction experiments, we conclude that LamB protein was integrated into, rather than translocated across, the vesicle membrane.

We consider it unlikely that integration of LamB protein occurred into 'zones of adhesion'²⁴⁻²⁶ which could occur as trace contaminants in our INV preparation²⁷. Post-translational recentrifugation of INVs in a sucrose gradient revealed a precise isopycnic overlap of the integrated LamB with the peak of INVs (data not shown) and not with denser fractions where zones of

adhesion would be expected (ref. 27). Therefore, it appears that LamB protein is integrated into an inner membrane domain in our *in vitro* experiments; by extension, we propose that *in vivo* LamB protein is initially integrated into the inner membrane at widely distributed translocation sites. Subsequent sorting to the outer membrane may proceed by a rapid migration to localized transport sites where a membrane fission/fusion process involving a vesicular shuttle could be used to deliver the protein to the outer membrane. Regions of such transport activity might appear as 'Bayer's junctions' or 'zones of adhesion' in electron micrographs²⁴⁻²⁷. Obviously, active LamB protein could not be tolerated in the cytoplasmic membrane of *E. coli* because of its function as a nonspecific diffusion channel⁸; however, LamB protein could be maintained in an inactive conformation by a variety of mechanisms, for example prevention of trimerization, since LamB protein monomers are probably inactive⁷.

It is not known whether the cytoplasmic membrane is the initial site of integration for all integral membrane proteins of the Gram-negative outer membrane or whether some proteins are first translocated into the periplasm followed by integration directly into the outer membrane²⁸.

To determine whether LamB protein could also be integrated into eukaryotic microsomal membranes, we translated LamB mRNA (and, as a control, also rabbit globin mRNAs) in a wheat germ cell-free translation system²⁹ supplemented with canine signal recognition particle (SRP) and/or salt-extracted canine microsomal membranes (KRM) which are depleted in SRP (refs 30-33) (Fig. 2). In the absence of both SRP and KRM, pre-LamB and globins are the major translation products (Fig. 2, lane 1). Minor polypeptides migrating faster than pre-LamB may have resulted from initiation at downstream methionines or from degradation of pre-LamB. Addition of SRP led to a considerable inhibition of synthesis of pre-LamB but not of the globins (Fig. 2, lane 2), indicating that SRP probably caused elongation arrest^{31,34}. Addition of SRP and KRM yielded the synthesis of three distinct polypeptides migrating slightly faster than pre-LamB (Fig. 2, lane 3). Post-translational incubation with proteinase K showed protection of all three polypeptides but not of pre-LamB and the globins (Fig. 2, lane 4). This protection was afforded by the microsomal membrane as solubilization with Triton X-100 yielded complete digestion (Fig. 2, lane 5). Addition of KRM only (Fig. 2, lane 6) yielded no protease-protected forms (lane 7).

Only the fastest moving of the three new polypeptides segregated by canine microsomal membranes migrated at a position of the mature form of LamB protein, suggesting that the two slower-migrating forms (in a position between pre-LamB and mature LamB) may represent mature, but modified forms of LamB protein. To determine whether they might represent glycosylated forms of mature LamB protein, they were incubated with endoglycosidase (Endo-H) (Fig. 2). Indeed, both the forms of intermediate relative molecular mass (M_r) (Fig. 2, lane 9 compared with 11), but not pre-LamB (lane 8 compared with 10), were converted after Endo-H treatment to a single polypeptide co-migrating with mature LamB protein. Inspection of the published amino-acid sequence of LamB protein¹⁶ indeed revealed two potential Asn-X-Thr glycosylation sites at residues 165-167 and 301-303. These observations suggested that the two forms represent mature LamB protein with one and two Asn-linked core oligosaccharides, although we cannot exclude the possibility that they result from differential trimming of a single Asn-linked core oligosaccharide.

The microsomal membrane-segregated forms of LamB protein, but not the globin products, sediment with the membranes at neutral pH (Fig. 3, lanes 2, 3) but are completely extracted at alkaline pH (lanes 4, 5). Because these forms were also protected from protease digestion (Fig. 2, lane 4), it can be concluded that they were translocated across rather than integrated into the microsomal membrane. To ensure that the observed alkaline extractability was not due to incomplete sedimentation of the

membranes, a derivative of glycosylated opsin which was known to be integrated into microsomal membranes³⁶ was shown to be alkaline inextractable under identical sedimentation conditions (Fig. 3, compare lanes 6-9).

Thus, nascent LamB protein has a distinctly different fate in the *E. coli* INVs, where it is integrated into the membrane of the vesicles, than in eukaryotic microsomes, where it is translocated into the lumen of the vesicles. Relatively little is known of the detailed mechanism by which polytopic integral membrane proteins are inserted into the phospholipid bilayer, although two different models have been proposed^{1,37,38}. Several investigators have suggested that insertion could occur by spontaneous physical partitioning of hydrophobic protein sequences into the hydrophobic core of the lipid bilayer without the assistance of catalytic protein^{37,38}. To explain the data presented here on the basis of such a model, one would have to invoke the differences in the lipid composition of the *E. coli* cytoplasmic membrane³⁹ and the canine rough endoplasmic reticulum membrane⁴⁰.

An alternative model proposes that integration of polytopic proteins into appropriate acceptor membranes is mediated by a translocation apparatus which decodes alternating signal and stop-transfer sequences^{1,36}. In the context of this model, our experimental results would be explained as follows. The translocation apparatus of the *E. coli* INVs was able to decode both the signal sequence(s) and the stop-transfer sequence(s) of LamB protein, leading to integration of the protein into the membrane. In contrast, the corresponding apparatus in the microsomal membrane was able to interpret the LamB signal sequence but not its stop-transfer sequence, leading to complete translocation of the protein.

We thank Dr Spencer Benson and Dr Martin Friedlander for providing us with the SP64 plasmid containing the LamB gene and the SP65 plasmid containing the derivative of bovine opsin cDNA, respectively, and Ms Gisele Nimic for typing. This research was supported by NIH grant PHS GM-27155.

Received 12 May; accepted 24 June 1986

1. Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496-1500 (1980).
2. Nikaido, H. & Vaara, M. *Microbiol. Rev.* **49**, 1-32 (1985).
3. Rosenbusch, J. P. *J. biol. Chem.* **249**, 8019-8029 (1974).
4. Ishii, J. N., Okajima, Y. & Nakee, T. *FEBS Lett.* **134**, 217-220 (1981).
5. Palva, E. T. & Westermann, P. *FEBS Lett.* **99**, 77-80 (1979).
6. Nakae, T. & Ishii, J. N. *Annls Microbiol. Inst. Pasteur, Paris* **133A**, 21-25 (1982).
7. Dorset, D. L., Engel, A., Massalski, A. & Rosenbusch, J. P. *Biophys. J.* **45**, 128-129 (1984).
8. Boehler-Kohler, B. A., Boos, W., Dieterle, R. & Benz, R. *J. Bact.* **138**, 33-39 (1979).
9. Schindler, H. & Rosenbusch, J. P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2302-2306 (1981).
10. Luckey, M. & Nikaido, H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 167-171 (1980).
11. Ferenci, T. & Boos, W. *J. J. supramolec. Struct.* **13**, 101-116 (1980).
12. Neuhaus, J. M., Schindler, H. & Rosenbusch, J. P. *EMBO J.* **2**, 1987-1991 (1983).
13. Wandersman, C., Schwartz, M. & Ferenci, T. *J. Bact.* **140**, 1-13 (1977).
14. Bavoil, P. & Nikaido, H. *J. biol. Chem.* **256**, 11385-11388 (1981).
15. Randall-Hazelbauer, L. & Schwartz, M. *J. Bact.* **116**, 1436-1446 (1973).
16. Clement, J. M. & Hofnung, M. *Cell* **27**, 507-514 (1981).
17. Müller, M. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7421-7425 (1984).
18. Marchal, C., Perrin, D., Hedgpeth, J. & Hofnung, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1491-1495 (1980).
19. Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057-7070 (1984).
20. Schnaitman, C. A. *J. Bact.* **104**, 890-901 (1970).
21. Steck, T. L. & Yu, J. *J. supramolec. Struct.* **1**, 220-248 (1973).
22. Fujiki, Y., Hubbard, A. L., Fowler, S. & Lazarow, P. B. *J. Cell Biol.* **93**, 97-102 (1982).
23. Gilmore, R. & Blobel, G. *Cell* **42**, 497-505 (1985).
24. Smit, J. & Nikaido, H. *J. Bact.* **135**, 687-702 (1978).
25. Bayer, M. H. in *Membrane Biogenesis* (ed. Tzagoloff, A.) 383-437 (Plenum, New York, 1975).
26. Bayer, M. H., Costello, G. P. & Bayer, M. E. *J. Bact.* **149**, 758-767 (1982).
27. Ishidate, K. *et al. J. biol. Chem.* **261**, 428-443 (1986).
28. Pugsley, A. P. & Schwartz, M. *FEMS Microbiol. Rev.* **32**, 3-38 (1985).
29. Erickson, A. H. & Blobel, G. *Meth. Enzym.* **96**, 38-49 (1983).
30. Walter, P. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7112-7116 (1980).
31. Müller, M., Ibrahim, I., Chang, C. N., Walter, P. & Blobel, G. *J. biol. Chem.* **257**, 11860-11863 (1982).
32. Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 682-691 (1983).
33. Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 84-93 (1983).
34. Walter, P. & Blobel, G. *J. Cell Biol.* **91**, 557-561 (1981).
35. Erickson, A. H. & Blobel, G. *J. biol. Chem.* **254**, 11771-11774 (1979).
36. Friedlander, M. & Blobel, G. *Nature* **318**, 338-343 (1985).
37. Van Heijne, G. & Blomberg, C. *Eur. J. Biochem.* **97**, 175-181 (1979).
38. Engelman, D. M. & Steitz, T. A. *Cell* **23**, 411-422 (1981).
39. Mizushima, S. & Yamada, H. *Biochim. biophys. Acta* **375**, 44-53 (1975).
40. Dewey, M. M. & Barr, B. *Curr. Topics Membrane Transport* **1**, 6 (1970).

Striking similarities between antigen receptor J pieces and sequence in the second chain of the murine CD8 antigen

Pauline Johnson & Alan F. Williams

MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

The CD8 antigen is a marker for T-lymphocyte subsets that is absent from helper T cells but expressed on cytotoxic T cells which recognize foreign determinants in association with class I major histocompatibility complex (MHC) antigens¹. It has been suggested that CD8 plays some part in recognition by CD8⁺ cytotoxic T cells since anti-CD8 antibodies can block their functions² and the human CD8 antigen contains a domain with clear similarities to immunoglobulin and T-cell receptor (TCR) variable-region (V) domains^{3,4}. Human CD8 antigen is thought to be a homodimer⁵ but in the mouse and rat the equivalent antigens (alternatively called Lyt2,3 and OX8) are heterodimeric^{1,6}. Rat CD8 contains two chains of relative molecular mass 32,000 (32K) and 37K: the 32K chain is the rat homologue of human CD8 and mouse Lyt2 (ref. 6). We describe here the molecular cloning of the rat 37K chain using an oligonucleotide probe predicted from peptide sequence. The full protein sequence is derived from the complementary DNA and matches limited peptide sequence for mouse Lyt3. The new sequence is more like immunoglobulin and T-cell receptor V domains than other T-cell antigens and includes a patch that is almost identical to some joining (J) piece sequences. This suggests that the CD8 and receptor heterodimers may have evolved directly from a common ancestor.

In previous studies the rat CD8 (OX8) antigen was purified from thymocytes, the chains separated and a limited amount of peptide sequence obtained from each chain⁶. From the 37K chain two peptides were isolated with the sequences ELQDSNKNK and FVKQF-K. A 14-residue oligonucleotide probe with all possible coding combinations was predicted from the amino-acid sequence QDSNK. To keep the variations to the minimum of 48 the probe was synthesized as two mixtures complementary to the coding sequence and then ³²P-labelled and used to screen rat thymocyte cDNA libraries (see Fig. 1 legend). Positive clones were mapped with restriction enzymes (Fig. 1a) and sequenced to give a predicted amino-acid sequence of 187 residues for the mature protein (Fig. 2). This sequence codes for the 37K chain because the two peptides previously isolated can be found at residues 36–44 and 181–187 respectively in the translated sequence. One discrepancy was found in that residue 2 in the FVKQF-K sequence was predicted as Met in the cDNA and Val in the peptide sequence. The phenylthiohydantoin derivatives of Val and Met elute at a very similar position on HPLC and re-examination of the trace showed that the residue originally designated Val was more likely to be Met. The FVKQF-K peptide also gave a double sequence in the protein sequencing⁶ which can now be interpreted as being due to polymorphism (see Fig. 2 legend).

The N-terminal residue of the translated 37K sequence is unknown but comparison with the N-terminal protein sequence of mouse Lyt3 (ref. 7) identified Leu as a likely first residue, as well as establishing that the rat 37K chain and Lyt3 are equivalent molecules in the two species.

Mouse Lyt3 sequence	L I Q A P S S L L V Q K
Rat 37K sequence	L L Q T P S S L L V Q T

A suitable initiator methionine residue is found 21 amino acids prior to the Leu residue assigned as position 1 (Fig. 2).

In Northern blot analysis a band of 1.3–1.6 kilobases (kb) was consistently found from rat thymocytes and T lymphocytes,

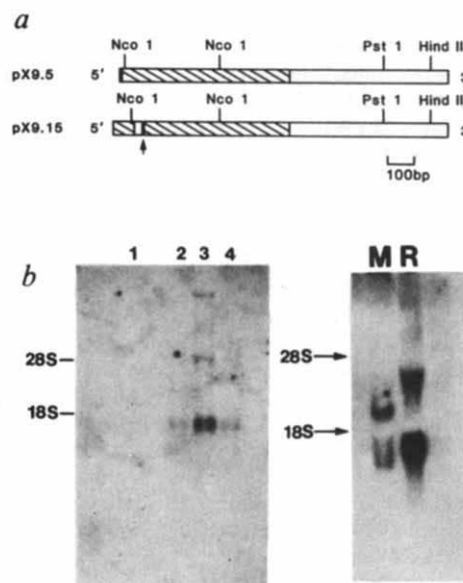


Fig. 1 *a*, Restriction-enzyme map of cDNA clones coding for the 37K chain. The hatched regions represent the signal sequence and coding region of the clones. Clone pX9.5 was from a rat AO strain thymocyte library (A. N. Barclay, unpublished) prepared by the RNase H method¹⁹, whilst pX9.15 was from a Sprague-Dawley thymocyte library²⁰ prepared by the loop-back method in which cDNA is commonly rearranged. The rearrangement in pX9.15 is shown by the arrow. The oligonucleotide probe was prepared as previously described²⁰ and the libraries were screened on nitrocellulose in 6×NET (1 M NaCl, 90 mM Tris-HCl pH7.4, 6 mM EDTA), 5×Denhardt's, 5% dextran sulphate, 0.5% SDS and 100 µg ml⁻¹ of denatured and sonicated herring sperm DNA at 30 °C, which is 4 °C below T_{hyb}, calculated for the lowest GC combination from 4(G+C)+2(A+T) (ref. 21). The filters were washed three times in 6×SSC, 0.1% SDS at 34 °C. *b*, Northern blot analysis using the rat 37K cDNA. Tracks 1–4 show hybridization with 5 µg RNA from rat B lymphocytes, T lymphocytes, cortical thymocytes and thymocytes respectively. M and R show hybridization with 50 µg of mouse and rat thymocyte RNA respectively. Total RNA^{20,22} was fractionated on a 1.4% (left) or 1.5% (right) agarose/formaldehyde denaturing gel, blotted on Gene-Screen (NEN) and hybridized with nick-translated pX9.15 as described in the GeneScreen manual. The blot on the left was washed with 2×SSC at 65 °C, whilst washing for that on the right was washed with 3×SSC at 65 °C.

while RNA from rat B lymphocytes gave no reaction (Fig. 1b). A band at about 3.5 kb was also seen which varied in intensity in different preparations. Analysis of mouse thymocyte RNA was done using low stringency conditions to reveal bands at 1.3 and 2.3 kb. The 3.5-kb (rat) and 2.3-kb (mouse) bands may be splicing intermediates and similar bands have been seen in Northern blot analysis with mouse Lyt2 cDNA⁸.

The predicted molecular architecture of rat CD8 is shown in Fig. 3. Both chains have V-like domains at the N-terminus (see below) followed by a short intervening sequence before a hydrophobic transmembrane sequence and a short cytoplasmic tail. The behaviour of the CD8 chains on SDS-polyacrylamide gel electrophoresis must be greatly influenced by the carbohydrate structures since the rat 32K chain, with one N-linked site, is equivalent to the mouse Lyt2 38K chain, with three N-linked sites^{1,8,9}, whilst the rat 37K chain with three sites is the equivalent of the mouse Lyt3 30K chain with one site^{1,10}. There is no known human equivalent of the rat 37K chain but this can now be further investigated using the rat cDNA in cross-hybridization studies.

The 32K and 37K rat CD8 chains show similarities throughout their sequences but only at the level of 21% identity (Fig. 4a).

Fig. 2 Nucleotide and corresponding amino-acid sequence of thymocyte cDNA clones coding for the rat 37K chain. cDNA clones described in Fig. 1a were digested with *AluI*, *HaeIII* or *NcoI* plus *PstI* restriction enzymes and cloned into M13mp8 or M13mp9 for sequencing by the dideoxy method, as used in ref. 20. The open reading frame is identified as coding sequence by comparison with peptide and mouse *Lyt3* sequences (see text). All the cDNA sequence on the 3' side of the stop codon was determined to yield 585 nucleotides without detecting a poly(A) addition site. The arrows and letters show the beginning of the V-like domain (V), J-related segment (J), membrane-proximal region (MP), transmembrane region (TM) and cytoplasmic domain (CY). The positions of three possible N-linked carbohydrate sites are also indicated. Sequencing of these clones plus another gave insight into a double sequence previously obtained for one of the peptides⁶ and interpreted as

FVKQF-K
-- FVKQF

had been produced by trypsin digestion of the succinylated 37K chain in which cleavage should occur only after Arg residues. Residues in the relevant region shown above are RIHFMKQFHK which is compatible with the minor sequence, allowing for the incorrect assignment of Val for Met (see text) and for the failure to detect Ile and His for technical reasons. In another clone, the relevant sequence was RIRFMKQFHK which would give rise to the major sequence. Thus the double-peptide sequence can be explained by polymorphism in this region and this is reasonable since the Sprague-Dawley rats, from which the protein and cDNA clones were isolated, are not inbred. One cDNA derived from AO strain rats gave the major sequence in this region.

GGAGCCCTGAAGGAGACAGGCTGCTCTCAAGAGCCCAAG	Met Gln Pro Trp Leu Trp Leu Val Phe Ser Val Lys Leu Ser	-8
ATG CAG CCA TGG CTC TGG CTG GTC TTC AGT GTG AAG CTA TCA		81
Ala Thr Gly Ser Ser Ala Leu Leu Gln Thr Pro Ser Ser Leu Leu Val Gln Thr Thr Ala Lys	→ V	17
GCT CTC TGG GGC AGC TCT GGC CTC CTT CAG ACT CCT TCA TCC CTG CTG GTT CAA ACC		153
Met Ser Cys Glu Ala Lys Thr Phe Pro Lys Gly Thr Thr Ile Tyr Trp Leu Arg Glu Leu Gln Asp Ser Asn		41
ATG TCC TGT GAG GCT AAA ACC TTC CCT AAA GGA ACA ACC ATC TAC TGG CTG AGA GAG CTC CAG GAT TCC AAC		225
Lys Asn Lys His Phe Glu Phe Leu Ala Ser Arg Thr Ser Thr Lys Gly Ile Lys Tyr Gly Glu Arg Val Lys		85
AAG AAC AAG GAC TTT GAG TTC CTG GCG TCC CGG ACT TCA ACC AAA GGA ATT AAA TAT GGT GAA AGG GTG AAG		297
Lys Asn Met Thr Leu Ser Phe Asn Ser Thr Leu Pro Phe Leu Lys Ile Met Asp Val Lys Pro Glu Asp Ser	→ J	89
AAG AAT ATG ACT CTG TCT TAT AAT TCA ACC CTG CCC TTT CTC AAG ATC ATG GAT GTG AAG CCA GAG GAC AGT		369
Gly Phe Tyr Phe Cys Ala Met Val Gly Ser Pro Met Val Val Phe Gly Thr Gly Thr Lys Leu Thr Val Val		113
GGC TTC TAC TTC TCG GCG ATG GTT GGG AGT CCC ATG GTA GTC TTT GGG ACA GGG ACA AAG CTG ATG GTG GTT	→ MP	441
Asp Val Leu Pro Thr Thr Ala Pro Thr Lys Lys Thr Thr Leu Lys Lys Lys Gln Cys Pro Thr Pro His Pro		137
GAT GTC CTT CTT CCA ACT GCT ACA ACC AAA AAG ACT ACA CTG AAG AAG AAA CAA TGT CCA ACC CCC CAC CCA		513
Lys Thr Gln Lys Gly Leu Thr Cys Gly Leu Ile Thr Leu Ser Leu Leu Val Ala Cys Ile Leu Val Leu Leu	→ TM	181
AAG ACC CAG AAA GGC TGT ACA TGT GGC CTT ATT ACC CTC ACC CTG CTG GTG GCG TGC ATC CTG GTT CTG CTG		585
Val Ser Leu Ser Val Ala Ile His Phe His Cys Met Arg Arg Arg Ala Arg Ile His Phe Met Lys Gln Phe	→ CY	185
GTA TCC CTC AGT GTG GCG ATC CAC TTT CAC TGC ATG CCG AGG AGA GCG CCG ATT CAC TTC ATG AAA CAG TTT		657
His Lys		187
CAC AAA TGA TCCGCCGCCACGACACTGACAGCCTGCTCAAGAGGACTCAATATAAGCAAGACCTGGACCAAGAGAGAGGACGATTTAGC		749
TCAGCTGTCTGCTGAAGCTGCTGCAAGTCTGTAATGGTCAGGACGTTACCTGGAATCTCAGAGAGAGGCCCAAGATCCAGGCTCACTCA		844
CGAGATGTGCTGAAGAACCAAGAACTCCACCATGACACCCCTGTGGCTTTTGATCCCTCTTGAGCTGGACCTTGGATGGCAGCCTTGACGCCA		939
CCATCTTTGTAAGTGTCTGTAAGGAGCAGCATGCTGCTTTTGGGCTCATGGAGAGAGCTAGGTGGCAGCCCTGCAGAGACTGCACATC		1034
CTCACACAGAGAGGCCACCTCATGAATAAGTTTCTCTGCAAAACATCTCTGGGCCCTTGGCAGCATCTGCTCTCAGGCCCTGCTGTGTTT		1129
TGGAGAGCTGTGCTGAGAGAGAGACTTCTACTTCATCTCTAGTAGGAATTTCTTAAGCTCTCTCCCTGCTGGCTTATATATTCTTTC		1224
CTCACTCTGTGCTTTTATGAATAATGCTCTGGATC		1261

In fact in the N-terminal domain of the sequences the 37K chain gives more identities with immunoglobulin and TCR V domains than with the 32K CD8 chain. For the V domains shown in Fig. 4a the number of identities with the first 113 residues of the 37K chain are: V_H, 33; V_λ, 30; V_κ, 30; TCR α, 32; TCR β, 27. In Fig. 4b the comparisons are quantitated by standard deviation scores obtained in the ALIGN program of Dayhoff and colleagues¹¹ and these data support the view that the 37K chain fits better with the V domains than do other immunoglobulin-related sequences.

This close similarity with receptor V domains is also indicated by the fact that the 37K sequence includes a patch that is almost identical to some V-domain J-piece sequences (Fig. 4c). The best matches are seen with J sequences from immunoglobulin V_λ, where up to 10 out of 11 identities are obtained, and mouse TCR α, clone TA39, which has a patch of 9 out of 10 identities. Patches with similarities to J pieces have also been identified in the sequences of the T-cell antigens CD4 (ref. 12) and MRC OX2 (ref. 13) but the matches with the 37K chain are much more impressive than for these other antigens.

A key point of interest is whether or not the sequence that matches the J pieces is separated from the rest of the V domain in the germ line as is the case for immunoglobulin and TCR V domains. To study this, a λ-phage genomic clone was isolated from a rat liver DNA library¹⁴, using a probe obtained after *NcoI* digestion of cDNA clone pX9.5 which included most of the 37K V-like domain (Fig. 1a). *EcoRI* digestion of the isolated λ clone gave a fragment of 2.8 kb that hybridized with the V-domain probe and this was digested with *NcoI* which should cleave on the 5' side of the J-like piece (nucleotides 402-407, Fig. 2). Sequencing in both directions from this site showed that there was no intervening sequence between the V-like domain and the J-like sequence. The exon was found to end at nucleotide 439 (amino acid 113).

The sequencing of both chains of the rat CD8 antigen reveals a structure that looks remarkably like an antigen receptor in that it consists of a heterodimer of dissimilar V-like domains. The existence of a sequence in the 37K chain that matches closely with that of V-domain J pieces supports the idea of a

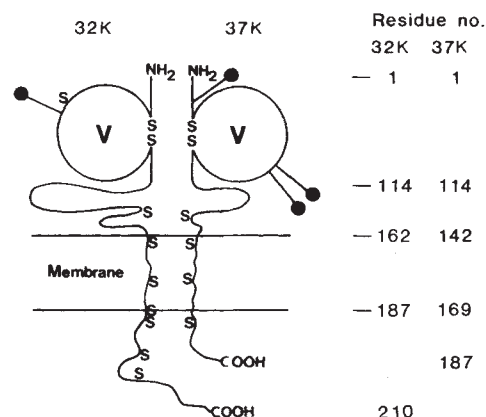


Fig. 3 Schematic diagram showing both chains of the rat CD8 antigen. Cysteine residues are indicated with an S and possible N-linked glycosylation sites by N. The V indicates that the segment shown as a circle is likely to have a fold like an immunoglobulin V domain and a disulphide between the relevant Cys residues is postulated. There must be at least one interchain disulphide bond and this may be in the membrane proximal region.

close relationship to receptor V domains. Genes coding for mouse *Lyt2* and *Lyt3* chains and for human CD8 are closely linked to the loci for immunoglobulin κ-chains^{15,16}. It has been postulated that CD8 functions to interact with class I MHC antigens on target cells^{17,18} and the structure would certainly fit with this idea, given the similarities between CD8 and receptors and the fact that the TCR recognizes foreign determinants in association with MHC antigens. It seems possible that CD8 and the receptors might have arisen from a common immediate precursor with a heterodimer structure like TCR. In the evolution of CD8 this structure may have been simplified by the loss of constant-region (C)-like domains in each chain, while the production of antigen receptors required the evolution of sets of V domains and of intervening sequences between the V domains and J segments. The splicing out of these intervening sequences



Fig. 4 Similarities between the 37K chain and immunoglobulin and TCR V-domain sequences. **a**, Alignment of the CD8 37K chain with the CD8 32K chain over their entire sequences and between the first 113 residues of the 37K chain and V domains from immunoglobulin and TCR chains. The dashes indicate gaps to maximize the alignments and boxes the positions where there are identities with the 37K residues. The bars above the sequences indicate positions of β -strands in the immunoglobulin fold, as determined from the structure of V_{λ} (ref. 23). The V domains are mouse V_H M36-65 (ref. 24); human V_{λ} , HA subgroup 1 and mouse V_{κ} , TEPC 111 (ref. 25); mouse TCR V_{β} , 86 Ti (ref. 26). **b**, Sequence comparisons using the ALIGN program¹¹ between rat and human CD8 sequences and V domains from immunoglobulin and TCR chains, as shown above, plus V-like sequences from Thy-1 (ref. 25) and MRC OX2 (ref. 28). The ALIGN program was run using the Mutation Data Matrix (250 PAMs) + 2 and a break penalty of 8. The best alignment is determined by the program and the score for this expressed as s.d. away from the mean score obtained from 140 random alignments. Scores of 3, 5.5 and 8 s.d. units indicate a probability of the relationship being due to chance of 0.135×10^{-2} , 0.19×10^{-7} and 0.11×10^{-18} (ref. 11). **c**, Alignments between residues 102-114 of the 37K sequence with J-piece sequences from immunoglobulin and TCR V domains and J-like sequences from CD4 and MRC OX2 antigens^{12,13}. Residues that are identical to any of the 37K residues are boxed and dashes indicate gaps to maximize the similarities. The J-piece sequences for immunoglobulin chains were determined from ref. 25 while those for TCR V_{β} and TCR V_{α} clone TA39 are taken from refs 29 and 30.

in differentiation may have first evolved as part of a superior genetic mechanism for somatic commitment to various single V-domain genes from among the multiple loci in the genome.

We thank Keith Gould for help with the oligonucleotide synthesis, Dr Neil Barclay for cDNA libraries, Dr George Gutman for the rat genomic DNA library, Protein Identification Resource for the NBRF database and the ALIGN program and Caroline Griffin for manuscript preparation. P.J. is a Beit Memorial Fellow.

Received 16 May; accepted 19 June 1986.

1. Walker, I. D. *et al. Immun. Rev.* **82**, 44-77 (1984).
2. MacDonald, H. R. *et al. Immun. Rev.* **68**, 89-115 (1982).
3. Littman, D. R. *et al. Cell* **40**, 237-246 (1985).
4. Sukhatme, V. P. *et al. Cell* **40**, 591-597 (1985).
5. Snow, P. M. & Terhorst, C. *J. biol. Chem.* **258**, 14675-14681 (1983).
6. Johnson, P., Gagnon, J., Barclay, A. N. & Williams, A. F. *EMBO J.* **4**, 2539-2545 (1985).
7. Walker, I. D. *et al. Immunogenetics* **23**, 60-63 (1986).
8. Zamoyska, R. *et al. Cell* **43**, 153-163 (1985).
9. Nakauchi, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5126-5130 (1985).
10. Luscher, B. *et al. J. Immun.* **135**, 1937-1944 (1985).
11. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. *Meth. Enzym.* **91**, 524-545 (1983).
12. Maddon, P. J. *et al. Cell* **42**, 93-104 (1985).
13. Barclay, A. N., Clark, M. J. & McCaughan, G. W. *Biochem. Soc. Symp.* **51**, 149-157 (1986).
14. Sheppard, H. W. & Gutman, G. A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7064-7068 (1981).
15. Gottlieb, P. D. *J. exp. Med.* **140**, 1432-1437 (1974).
16. Sukhatme, V. P. *et al. J. exp. Med.* **161**, 429-434 (1985).
17. Reinherz, E. L., Meurer, S. C. & Schlossman, S. F. *Immun. Today* **4**, 5-8 (1983).
18. Swain, S. L. *Immun. Rev.* **74**, 129-142 (1983).
19. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
20. Thomas, M. L., Barclay, A. N., Gagnon, J. & Williams, A. F. *Cell* **41**, 83-93 (1985).
21. Wallace, R. B. *et al. Nucleic Acids Res.* **9**, 897-894 (1981).
22. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
23. Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48**, 961-997 (1979).
24. Siekevitz, M. *et al. Eur. J. Immun.* **12**, 1023-1032 (1982).
25. Kabat, E. A. *et al. Sequences of Proteins of Immunological Interest* (US Department of Health and Human Services, NIH, Bethesda, 1983).
26. Hedrick, S. M. *et al. Nature* **308**, 153-158 (1984).
27. Chien, Y.-h. *et al. Nature* **312**, 31-35 (1984).
28. Clark, M. J., Gagnon, J., Williams, A. F. & Barclay, A. N. *EMBO J.* **4**, 113-118 (1985).
29. Chien, Y.-h. *et al. Nature* **309**, 322-326 (1984).
30. Arden, B., Klotz, J. L., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).

Transcripts encoded by a homoeo box gene are restricted to dorsal tissues of *Drosophila* embryos

Helen J. Doyle, Katherine Harding, Timothy Hoey & Michael Levine

Department of Biological Sciences, Columbia University, New York, New York 10027, USA

In *Drosophila* elaboration of positional identity along the anterior-posterior and dorsal-ventral embryonic body axes involves early zygotic gene functions that are expressed in response to maternal cues present in the unfertilized egg¹. Zygotic loci that are required for the specification of positional identity along the anterior-posterior body axis have been described in detail²⁻⁹. Less is known about the zygotic loci responsible for differentiation of the dorsal-ventral pattern; however, several genes that might be involved have been identified^{3-5,10}. *Zerknullt* (*zen*) is an example of a zygotic gene required for correct differentiation of dorsally derived embryonic tissues¹⁰. On the basis of homoeo box cross homology, we have now isolated a gene, called *S60*, that derives from the *zen* region of the Antennapedia complex (ANT-C)^{7,10}. Transcripts encoded by *S60* transiently accumulate in the dorsal-most tissues of developing embryos. This pattern of expression suggests that *S60* corresponds to *zen*. Since *S60* contains a homoeo box, it is possible that differentiation of the anterior-posterior and dorsal-ventral embryonic patterns involves similar molecular mechanisms.

There are at least 20 homoeo box genes in *Drosophila* (W. McGinnis, personal communication). Thus far, 11 of these genes have been isolated and characterized¹¹⁻²³. Using the homoeo

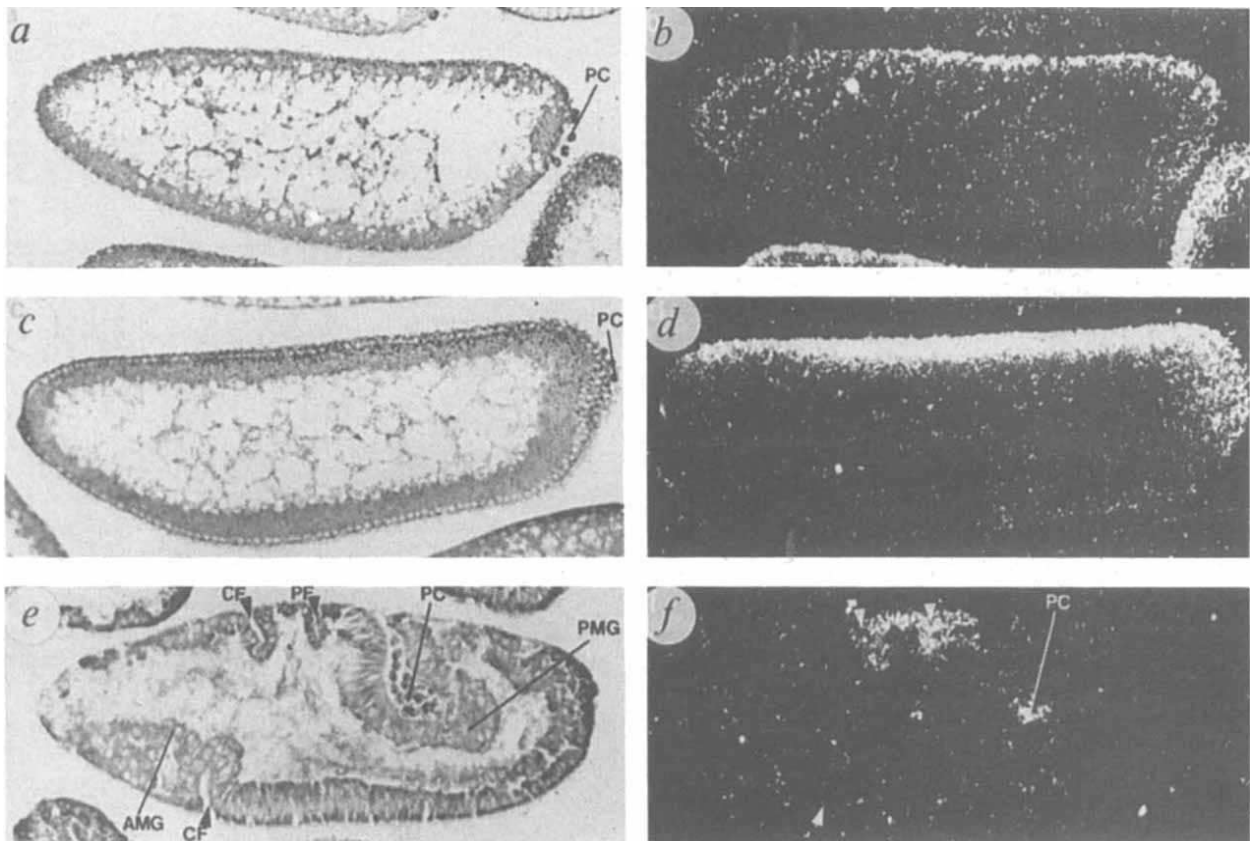


Fig. 3 Distribution of *S60* transcripts in 2-5-h wild-type embryos. All sections are sagittal and oriented such that anterior is to the left and dorsal is up. The tissue autoradiograms were hybridized with an *S60* cDNA probe. *a*, Localization of *S60* transcripts in a cleavage stage 11-12 embryo, soon after formation of the syncytial blastoderm. Specific hybridization signals are detected over the dorsal surface and posterior pole. The labelling does not include the pole cells. *b*, Dark-field photomicrograph of *a*. *c*, *S60* transcripts in a cleavage stage 14 embryo, just before cellularization. Strong hybridization signals are detected along the length of the dorsal surface and over the somatic nuclei of the posterior pole. *d*, Dark-field photomicrograph of *c*. *e*, Embryo undergoing germ band elongation ($\sim 4\frac{1}{2}$ -5-h post-fertilization). *S60* hybridization signals are localized to the presumptive amnioserosa (bracketed by the arrowheads) and to a subpopulation of the pole cells within the posterior midgut invagination. *f*, Dark-field photomicrograph of *e*. Abbreviations: AMG, anterior midgut invagination; CF, cephalic furrow; PC, pole cells; PMG, posterior midgut invagination; PF, posterior transverse furrow.

Methods. Paraffin sections of *Drosophila* embryos were prepared as described in ref. 39. The tissues were treated and hybridized with the *S60* cDNA probe after ^3H -labelling by nick translation⁴⁰.

begins, *S60* transcripts are restricted to a dorsal strip 5-6 cells wide that extends along the length of the embryo. Figure 4*a, b* shows a cross-section through an early cleavage stage 14 embryo after hybridization with the *S60* probe. *S60* transcripts encompass about 30% of the embryo's circumference. By the time gastrulation begins, hybridization signals are confined to the dorsal-most region of the embryo, encompassing ~ 5 -6 cells (less than 10% of the embryo's circumference) (Fig. 4*c, d*). A single strip of *S60* hybridization is observed in a horizontal section through the dorsal surface of a gastrulating embryo (Fig. 4*e, f*).

Thus *zen* appears to be one of the zygotic genes required for differentiation of the dorsal-ventral pattern (C. Nüsslein-Volhard, personal communication)¹⁰. In *zen* mutant embryos, several dorsally derived tissues, including the amnioserosa, fail to form. The absence of these dorsally derived tissues leads to a failure of germ band elongation, which results in gross disruptions of the germ band. Many of the defects seen in older *zen*⁻ embryos, such as failure of head involution, may arise as secondary consequences of these early disruptions¹⁰. The temporal and spatial pattern of *S60* expression correlates well with the time and location of *zen* gene function.

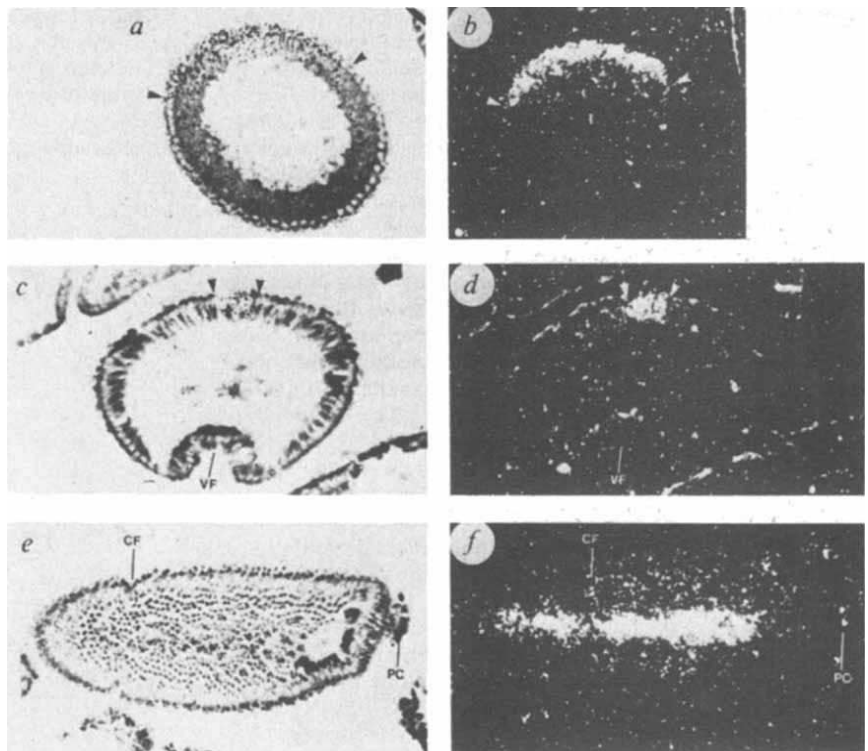
With the exception of *S60*, each of the known homoeo box genes appears to be involved with the elaboration of positional identity along the anterior-posterior embryonic body axis (reviewed in ref. 20). It has been proposed that the *Drosophila*

homoeo box is generally involved in the specification of meta-mers during embryogenesis²⁵. However, *S60* contains a well-conserved copy of the homoeo box and appears to affect an embryonic process distinct from those specified by other homoeo box genes. This suggests that the common denominator of homoeo box gene function does not involve segmentation *per se*, but instead involves a more fundamental embryonic process, such as the specification of positional identity.

Since *S60* contains a well-conserved homoeo box, it is possible that the *S60* protein functions in a way similar to other homoeo box gene products. Direct and/or indirect regulatory interactions among homoeo box genes expressed along the anterior-posterior body axis have been shown previously^{22,31-34}. In particular, it appears that the maintenance of the specific expression patterns of several segmentation and homoeotic genes is dependent on the activities of other homoeo box genes. Although *S60* is expressed in a pattern different from other known homoeo box genes, it is possible that the *S60* protein interacts with anterior-posterior homoeo box gene products. Thus, the conservation of the homoeo box may provide a molecular link between the formation of the anterior-posterior and dorsal-ventral patterns.

Ten maternally-expressed loci have been shown to be required for the establishment of the basic features of the dorsal-ventral pattern^{1,35,36}. Embryos that completely lack a product encoded by any one of these loci show altered identities of ventral cells

Fig. 4 *S60* transcripts are confined to the dorsal-most regions of the ectoderm. All tissue autoradiograms were hybridized with an *S60* cDNA probe. *a-d*, Cross-sections are oriented so that dorsal is up. The photomicrographs in *e* and *f* show a horizontal section through the dorsal surface of a gastrulating embryo; anterior is to the left. *a*, Cross-section through an early cleavage stage 14 embryo. *S60* transcripts are detected over ~30% of the embryo's circumference. *b*, Dark-field photomicrograph of *a*. *c*, Gastrulating embryo showing invagination of the ventral furrow. *S60* transcripts are localized to a strip 5-6 cells wide running the length of the embryo along the dorsal surface. *d*, Dark-field photomicrograph of *c*. *e*, Horizontal section through the dorsal ectoderm of a gastrulating embryo. *f*, Dark-field photomicrograph of *e*. The intensity of the signal seen in *f* is weaker toward the anterior end, and there appear to be two gaps in the labelling. The arrowheads indicate the limits of the hybridization signal. Abbreviations: CF, cephalic furrow; PC, pole cells; VF, ventral furrow. For methods see legend to Fig. 3.



such that they acquire a more dorsal state of differentiation. As a result these mutants show defects in gastrulation and germ-band elongation. So far, no zygotically-expressed gene has been shown to be expressed in a dorsal-ventral specific pattern. It is possible that the initiation of *S60* expression and the evolution of its final pattern are in direct response to information encoded by one or more of these maternal genes. It is also likely that other zygotic genes will show distinctive dorsal-ventral patterns of expression. The isolation of maternal and other zygotic genes involved in dorsal-ventral pattern formation will help to elucidate the genetic and molecular interactions among these genes.

We thank Christiane Nüsslein-Volhard and Bill McGinnis for sharing unpublished results, and Hanna Radomska for technical assistance. This work was supported by grants from the NIH and the Searle Scholars Program.

Received 9 April; accepted 4 June 1986.

- Nüsslein-Volhard, C. in *Determinants of Spatial Organization* (eds Subtelney, S. & Konigberg, I. R.) 185-211 (Academic, New York, 1979).
- Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
- Nüsslein-Volhard, C., Wieschaus, E. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 267-282 (1984).
- Wieschaus, E., Nüsslein-Volhard, C. & Jürgens, G. *Wilhelm Roux Arch. dev. Biol.* **193**, 283-295 (1984).
- Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 296-307 (1984).
- Lewis, E. B. *Nature* **276**, 565-570 (1978).
- Kaufman, T. C., Lewis, R. & Wakimoto, B. *Genetics* **94**, 115-133 (1980).
- Wakimoto, B. T. & Kaufman, T. C. *Dev. Biol.* **81**, 51-64 (1981).
- Ouweneel, W. H. *Adv. Genet.* **18**, 179-205 (1976).
- Wakimoto, B. T., Turner, F. R. & Kaufman, T. C. *Dev. Biol.* **102**, 147-172 (1984).
- Bender, W. *et al. Science* **221**, 23-29 (1983).
- Garber, R. L., Kuroiwa, A. & Gehring, W. J. *EMBO J.* **2**, 2027-2036 (1983).
- Scott, M. P. *et al. Cell* **35**, 763-776 (1983).
- McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. & Gehring, W. J. *Nature* **308**, 428-433 (1984).
- Fjose, A., McGinnis, W. & Gehring, W. J. *Nature* **313**, 284-289 (1985).
- Karch, F. *et al. Cell* **43**, 81-96 (1985).
- Kuner, J. M. *et al. Cell* **42**, 309-316 (1985).
- Regulski, M. *et al. Cell* **43**, 71-80 (1985).
- Mlodzik, M., Fjose, A. & Gehring, W. J. *EMBO J.* **4**, 2961-2969 (1985).
- Hoey, T., Doyle, H. J., Harding, K., Wedeen, C. & Levine, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4809-4813 (1986).
- Levine, M. *et al. Cold Spring Harb. Symp. quant. Biol.* **50**, 209-222 (1985).
- Harding, K., Wedeen, C., McGinnis, W. & Levine, M. *Nature* **229**, 1236-1242 (1985).
- Kuroiwa, A., Kloter, U., Baumgartner, P. & Gehring, W. J. *EMBO J.* **4**, 3757-3764 (1985).
- Scott, M. P. & Weiner, A. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4115-4120 (1984).
- McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. *Cell* **37**, 403-408 (1984).

- Poole, S., Kauvar, L. M., Drees, B. & Kornberg, T. *Cell* **40**, 37-43 (1985).
- Zalokar, M. & Erk, F. J. *Microbiol. Cell* **25**, 97-106 (1976).
- Foe, V. & Alberts, B. J. *cell. Sci.* **61**, 31-70 (1983).
- Poulson, D. F. in *Biology of Drosophila* (ed. Demerec, M.) 168-274 (Wiley, New York, 1950).
- Campos-Ortega, J. A. & Hartenstein, V. *The Embryonic Development of Drosophila melanogaster* (Springer, Berlin, 1985).
- Hafen, E., Levine, M. & Gehring, W. J. *Nature* **307**, 287 (1984).
- Struhl, G. & White, R. A. H. *Cell* **43**, 507-519 (1985).
- Wedeen, C., Harding, K. & Levine, M. *Cell* **44**, 739-748 (1986).
- Howard, K. & Ingham, P. *Cell* **44**, 949-957 (1986).
- Anderson, K. V. & Nüsslein-Volhard, C. *Nature* **311**, 223-227 (1984).
- Anderson, K. V., Jürgens, G. & Nüsslein-Volhard, C. *Cell* **42**, 779-789 (1985).
- Lepesant, J. A., Keizlarova-Lepesant, J. & Garen, A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5570-5574 (1978).
- Rozeck, C. E. & Davidson, N. *Cell* **32**, 23-34 (1983).
- Ingham, P. W., Howard, K. R. & Ish-Horowitz, D. *Nature* **318**, 439-445 (1985).
- Hafen, E., Levine, M., Garber, R. L. & Gehring, W. J. *EMBO J.* **2**, 617-623 (1983).

Recombinant human TNF induces production of granulocyte-monocyte colony-stimulating factor

Reinhold Munker*, Judith Gasson*, Makio Ogawa† & H. Phillip Koeffler*‡

* Department of Medicine, Division of Hematology/Oncology, University of California, Los Angeles Medical Center, Los Angeles, California 90024, USA

† Department of Medicine, Veterans Administration Medical Center, Charleston, South Carolina 29403, USA

Tumour necrosis factor (TNF) is synthesized by macrophages exposed to endotoxin¹. It produces haemorrhagic necrosis of a variety of tumours in mice and is cytostatic or cytotoxic against various transformed cell lines *in vitro*, but viability of normal human or rodent cells is unaffected²⁻⁴. The role of TNF is unlikely to be restricted to the rejection of tumours. Colony-stimulating factors (CSFs) are required for survival, proliferation and differentiation of haematopoietic progenitor cells. The

‡ To whom reprint requests should be addressed.

haematopoietic growth factor known as granulocyte-monocyte colony-stimulating factor (GM-CSF) has the ability to stimulate proliferation and differentiation of normal granulocyte-monocyte and eosinophil stem cells and enhance the proliferation of pluripotent, megakaryocyte and erythroid stem cells⁵. In addition, GM-CSF stimulates a variety of functional activities in mature granulocytes and macrophages, for example inhibition of migration, phagocytosis of microbes, oxidative metabolism, and antibody-dependent cytotoxic killing of tumour cells⁶⁻⁷. We show here that TNF markedly stimulates production of GM-CSF messenger RNA and protein in normal human lung fibroblasts and vascular endothelial cells, and in cells of several malignant tissues.

Recombinant TNF stimulates normal human lung fibroblasts in a dose-response fashion to synthesize increasing amounts of CSF, with 25 U ml⁻¹ TNF stimulating 50% maximal production of CSF (Fig. 1a). CSF activity was determined by the ability of conditioned medium from the fibroblasts exposed to TNF to stimulate the clonal proliferation of granulocyte-monocyte stem cells (GM-CFC) from normal human marrow. A similar dose-response curve was observed when the same conditioned medium was tested for its ability to stimulate clonal proliferation of cells of a human leukaemic myeloblast line (KG-1) which is dependent on CSF for clonal growth (Fig. 1a)⁸. These results suggest that the growth factor present in the conditioned medium is having a direct effect on the clonogenic cells and not stimulating accessory cells to produce CSF. TNF alone had no effect on clonal growth of either GM-CFC or KG-1 cells. Time-response studies showed that CSF was synthesized within 1 day of exposure of fibroblasts to TNF and the maximal response occurred on day 3 (Fig. 1b).

The colonies that developed on day 10 of co-culture of normal human marrow cells with the conditioned medium from fibroblasts exposed to TNF (200 U ml⁻¹ for 3 days) were aspirated individually and analysed morphologically and histochemically⁹; a total of 41% granulocyte, 12% macrophage, 33% granulocyte/monocyte and 14% eosinophil colonies developed. The conditioned medium (5.0% (v/v) from 3 × 10⁵ fibroblasts exposed to 200 U ml⁻¹ TNF for 3 days) also stimulated production of multipotential and early erythroid human colonies. Umbilical cord blood was enriched for multipotential and erythroid precursors by panning with MY-10 monoclonal antibody and was cultured in methylcellulose in the presence of the conditioned medium and 2 U ml⁻¹ erythropoietin (method described in ref. 9). A mean of four multipotential and four erythroid burst colonies (BFU-E) formed per 3 × 10³ cells plated. Neither erythropoietin nor the conditioned medium alone stimulated either multipotential or BFU-E colonies. Taken together, these results suggest that GM-CSF is a growth factor being produced by the fibroblasts cultured with TNF⁵.

Further evidence was obtained using two heterologous rabbit antisera raised against purified biosynthetic (recombinant) GM-CSF (ref. 10). Antisera from both animals decreased the CSF activity of the conditioned medium by a mean 86% as compared with preimmune rabbit sera from the same animals (Table 1). A GM-CSF radioimmunoassay (RIA) showed that the fibroblasts (5 × 10⁵ ml⁻¹) cultured for 48 h with 0, 10, 100 and 1,000 U ml⁻¹ TNF synthesized undetectable, 2.9, 49.5 and 34 ng ml⁻¹ of GM-CSF, respectively. For comparison, the conditioned medium from mitogen-stimulated T lymphocytes (5 × 10⁵ ml⁻¹ exposed to 50 µg ml⁻¹ phytohaemagglutinin for 4 days), which is one of the richest known sources of GM-CSF from non-neoplastic human cells¹⁰, synthesized 5 ng ml⁻¹ of GM-CSF as determined by RIA.

Further studies examined the ability of a variety of normal and malignant human tissues to synthesize GM-CSF after exposure to 200 U ml⁻¹ TNF for 48 h (Table 1). Normal umbilical vein endothelial cells, smooth muscle cells from the aorta, lung squamous cancer and adenocarcinoma cells were stimulated in their synthesis of GM-CSF. In contrast, fibroblasts from skin and bone marrow, hepatoma cells, thymocytes and a

variety of lines of both B and T lymphocytes and myeloid haematopoietic cells were unable to synthesize detectable GM-CSF when cultured with TNF (Table 1).

These observations were confirmed by Northern analysis of mRNA. As shown in Fig. 1, normal lung fibroblasts contained undetectable mRNA for GM-CSF (lane 1). After 48 h exposure

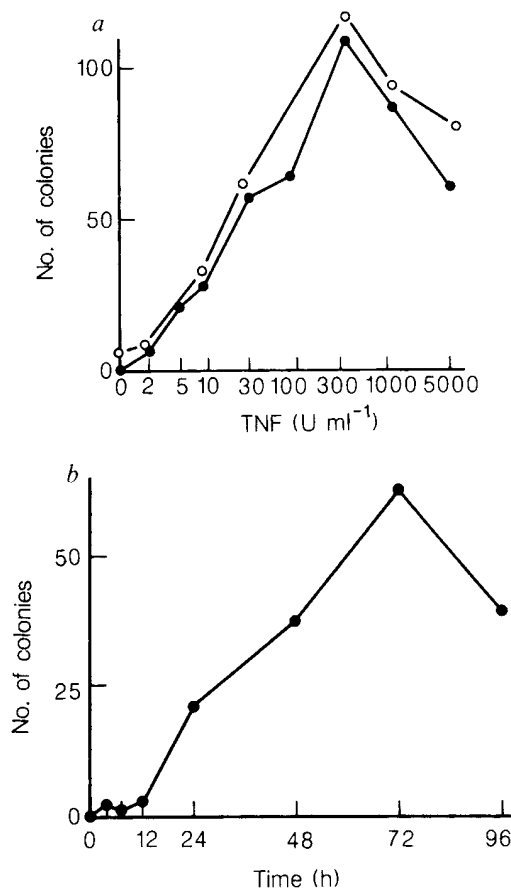


Fig. 1 Effect of TNF on CSF synthesis: dose- and time-response studies. Normal human pulmonary fibroblasts (1×10^5 ml⁻¹) were exposed to TNF at different concentrations (a) and for different lengths of time (b). Conditioned media were collected and tested for their ability to stimulate clonal proliferation of normal human marrow granulocyte-macrophage stem cells (GM-CFC, ●—●) or human KG-1 (ref. 8) myeloblastic leukaemia cells whose proliferation is dependent on CSF (○—○). Results represent means of two experiments with quadruplicate plates per point. TNF added directly to myeloid cells in soft agar did not stimulate clonal growth. **Methods.** The presence of CSF was determined by its ability to stimulate growth of myeloid colonies in agar. Normal bone marrow cells from adult volunteers were separated on Ficoll-Paque (1.077 g cm⁻³) (Pharmacia), and monocytes were depleted by 5 h adherence to 60-mm tissue culture dishes (Lux, Miles Laboratories) at 37 °C. The non-adherent bone marrow cells (1×10^5 ml⁻¹) or KG-1 cells (2×10^3 ml⁻¹) were plated in quadruplicate in 35-mm Petri dishes (Falcon) in a mixture containing 30% fetal calf serum (FCS) (Gibco), 1% bovine serum albumin (Sigma), 10^{-4} M mercaptoethanol (Sigma), 0.3% soft agar, and serial dilutions of the fibroblast conditioned medium or, as positive control, conditioned medium prepared from the human Mo T-lymphoblast cell line which is rich in CSF¹⁰. Plates were incubated at 37 °C in a high humidity, 5% CO₂-95% air incubator, and colonies were counted on day 14. There was no colony formation unless Mo or fibroblast conditioned media were added. Recombinant human TNF (gift of Genentech, lot 3238-14) was purified to >99% homogeneity with a final specific activity of 2.5×10^7 U per mg protein. The undiluted material contained <0.8 ng of endotoxin ml⁻¹ and this amount of lipopolysaccharide (endotoxin) had no effect when added to the colony-forming assays.

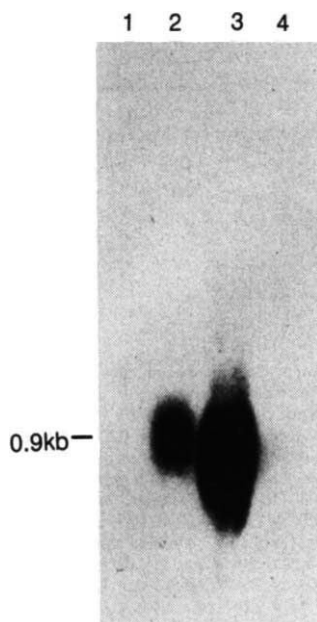


Fig. 2 Effect of TNF on GM-CSF mRNA synthesis by lung fibroblasts. Poly(A)-selected cytoplasmic RNA was prepared and analysed by agarose gel electrophoresis and nylon-base paper blotting as previously described¹⁷; the probe was gel purified full-length GM-CSF cDNA¹⁰ with specific activity of 2×10^8 c.p.m. per μg DNA. Lane 1, normal lung fibroblasts; lane 2, fibroblasts exposed to TNF (200 U ml^{-1}); lane 3, GM-CSF hyperproducer T-lymphocyte line infected with HTLV-1 (AB-VDR); lane 4, GM-CSF hyperproducer subline (1A-6) of bladder carcinoma line (5637) (ref. 18). AB-VDR cells ($1 \times 10^6 \text{ ml}^{-1}$) produced 823 ng ml^{-1} GM-CSF as determined by RIA which is nearly 160-fold greater production than mitogen-activated normal T lymphocytes ($1 \times 10^6 \text{ ml}^{-1}$). Exposure of autoradiogram >14-days continued to show no signal in lane 1 which contains fibroblasts not exposed to TNF. In contrast, all lanes showed comparable hybridization when the blot was reprobed with a β -actin cDNA (data not shown).

to 200 U ml^{-1} TNF, the fibroblasts contained appreciable mRNA for GM-CSF (lane 2). The 0.9-kilobase (Kb) hybridizing RNA is the same size as lymphocyte-derived GM-CSF mRNA¹⁰.

These data show that TNF modulates the synthesis of GM-CSF in several tissue types. Our findings have several major implications. First, GM-CSF will be synthesized locally at inflammatory sites by activation of tissue macrophages to produce TNF. This local GM-CSF could functionally activate mature white cells that migrate to inflammatory foci and, by inhibiting further migration, ensure their retention in the region of inflammation⁷. Second, the GM-CSF produced as a result of TNF synthesis by macrophages at the inflammatory site could enhance proliferation of haematopoietic stem cells. This would lead to increased production of the granulocytes and macrophages needed to fight infection. We and others¹⁹ have recently observed that recombinant interleukin-1 also stimulates GM-CSF synthesis of cultured lung fibroblasts and vascular endothelial cells in a similar fashion to TNF (manuscript in preparation).

Recently, other investigators showed that TNF specifically binds to many normal cell types¹¹, increases expression of HLA-A and -B antigens on vascular endothelial cells and dermal fibroblasts¹², enhances aggregation and adhesion of neutrophils to cultured endothelial cells^{13,14}, stimulates growth of normal fibroblasts^{15,16} and probably inhibits lipoprotein lipase activity of fat cells which may lead to cachexia during parasitic and bacterial infections¹¹. Our findings are consistent with TNF being a pleiotropic mediator of various activities in normal cells.

Table 1 GM-CSF production after exposure to tumour necrosis factor

Cell type (cell line)	TNF (200 U ml^{-1})	Normal human myeloid colony formation (mean)
Lung fibroblasts	0	0
	+	187
	+ and GM-CSF antibody C	22
	+ and GM-CSF antibody D	30
	+ and preimmune serum C	170
	+ and preimmune serum	164
Vascular endothelia	0	0
	+	155
Smooth muscles	+	18
	0	0
Lung squamous carcinoma (SKMES)	0	15
	+	74
Lung adenocarcinoma (SUT)	0	7
	+	65
Skin and bone marrow fibroblasts	+	0
Simian virus 40 infected fibroblasts (SV-80)	+	0
Thymocytes	+	0
Hepatoma cell line (Hep 2)	+	0
Myeloid leukaemia lines (K562, HL-60, U938)	+	0
Lymphoid leukaemia lines (NALM1, HCL, CEM)	+	0

The ability of TNF to stimulate production of CSF was determined as described in the legend to Fig. 1. TNF (200 U ml^{-1}) was co-cultured with 5×10^5 cells ml^{-1} of a variety of human cell types for 3 days in the presence of α medium plus 10% fetal calf serum. Normal umbilical vein endothelial cells were isolated by collagenase treatment of human umbilical cord veins, and cells were tested in their first passage; thymocytes were obtained at the time of cardiac surgery; smooth muscle cells were studied at their first passage after collection from aorta; NALM1 and HCL are B-lymphocyte lines and CEM is a T-cell line. The heterologous antisera raised in two rabbits (C and D) against purified biosynthetic (recombinant) GM-CSF¹⁰ and conditioned medium from lung fibroblasts exposed to 200 U ml^{-1} TNF were pre-cultured (v/v) for 1 h at 23°C and then added to culture dishes to test for CSF activity. Preimmune sera from both rabbits were also tested for anti-CSF activity.

We thank Suzanne Bookstaver and Regina Simon for secretarial help; Steven Clark (Genetics Institute, Boston, Massachusetts), Michael Shepard (Genentech, San Francisco), Jürgen Bessemer (Sandoz Institute), Peter Lomedico of (Hoffmann-La Roche) and Larry Souza (AmGen) for providing the GM-CSF and IL-2 probes, recombinant TNF, recombinant IL-2 protein, recombinant IL-1 protein and granulocyte-CSF respectively. H.P.K. is supported in part by NIH grants, the Dr Murray Geisler Memorial Leukemia Fund, the Louis Fagin Leukemia Research Foundation and an NIH Career Development Award. J.G. and H.P.K. are members of the Jonsson Comprehensive Cancer Center. R.M. is a Fellow of Deutsch Forschungsgemeinschaft.

Received 9 April; accepted 20 May 1986.

1. Carswell, E. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 3666-3670 (1975).
2. Helson, L., Green, J., Carswell, E. & Old, L. *Nature* **258**, 730-732 (1975).
3. Matthews, N., Watkins, J. F. *Br. J. Cancer* **38**, 302-309 (1978).
4. Aggarwal, B. *et al. Science* **260**, 2345-2354 (1985).
5. Metcalf, D. *Science* **229**, 16-22 (1984).

6. Gasson, J. *et al. Science* **226**, 1339-1342 (1984).
7. Weisbart, R., Golde, D., Clark, S., Wong, G. & Gasson, J. *Nature* **314**, 361-362 (1985).
8. Koefler, H. P. & Golde, D. *Science* **200**, 1153-1155 (1978).
9. Leary, A., Ogawa, M., Strauss, L. & Civin, C. *J. clin. Invest.* **74**, 2193-2197 (1984).
10. Wong, G. G. *et al. Science* **228**, 810-814 (1985).
11. Beutler, B., Mahoney, J., Le Trang, N., Pekala, P. & Cerami, A. *J. exp. Med.* **161**, 984-995 (1985).
12. Collins, T., La Pierre, L. A., Fiers, W., Strominger, J. & Pober, J. S. *Proc. natn. Acad. Sci. U.S.A.* **83**, 446-450 (1986).
13. Gamble, J. R., Harlan, J. M., Klebanoff, S. J., Vadas, M. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8667-8681 (1985).
14. Shalaby, M. R. *et al. J. Immun.* **135**, 2069-2073 (1985).
15. Sugarman, B. *et al. Science* **230**, 943-945 (1985).
16. Vilcek, J. *et al. J. exp. Med.* **163**, 632-643 (1986).
17. Maniatis, T., Fritsch, E. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
18. Souza, L. *et al. Science* **232**, 61 (1986).
19. Zucali, J. R. *et al. J. clin. Invest.* **77**, 1857 (1986).

Bovine chromogranin A sequence and distribution of its messenger RNA in endocrine tissues

Anna Iacangelo*, Hans-Urs Affolter*, Lee E. Eiden*‡, Edward Herbert† & Mark Grimes†

* Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, Maryland 20892, USA

† Institute for Advanced Biomedical Research, The Oregon Health Sciences University, Portland, Oregon 97201, USA

Chromogranin A is contained in storage vesicles of chromaffin cells of the adrenal medulla and released with catecholamines when the splanchnic nerve is stimulated^{1,2}. Chromogranin A is similar to secretory protein I (SP-I), a major secreted protein of the parathyroid³. Chromogranin A/SP-I immunoreactivity is abundant in endocrine cells that secrete peptide hormones from storage vesicles⁴⁻⁷. Chromogranins may act in neuroendocrine secretion by binding intravesicular calcium^{8,9}. Serum levels of chromogranin are raised in hypertension¹⁰ and endocrine neoplasia^{8,10,11}. We report here the isolation and sequencing of a cDNA encoding bovine chromogranin A, providing the first complete primary structure of a chromogranin protein. Chromogranin A is a highly acidic protein with an apparent relative molecular mass (M_r) of 75,000 on SDS-PAGE, but an actual M_r of 48,000. Adrenal medulla, brain, pituitary and parathyroid are all sites of synthesis of chromogranin A. The primary structure of chromogranin A, and the presence of chromogranin mRNA in the parathyroid, indicate that chromogranin A and SP-I are identical.

Bovine prechromogranin A was sequenced from its complementary DNA as described in Fig. 1. The longest clone obtained, pCHRG4A, contained an 1883 base insert with a single long open reading frame encoding a protein of M_r 50,000 (50K) (Fig. 1B). The size of this clone (allowing for a poly(A) tail of 200 nucleotides) is consistent with the 2,100-base messenger RNA (mRNA) detected by the cDNA in adrenal medulla (Fig. 4A). The amino-acid composition of the protein encoded by the cDNA clone is identical to that reported for chromogranin A and secretory protein I (ref. 3). Prechromogranin A contains no sites for N-linked glycosylation, consistent with the presence of O-linked sugars in the native glycoprotein¹². The deduced N-terminal amino-acid sequence of chromogranin A is identical to that reported for human and bovine chromogranin A by Kruggel *et al.*¹³, and differs by two amino-acids from that reported by Angeletti¹⁴ and Settlemann *et al.*¹⁵ for bovine chromogranin A. The latter may represent the N-terminal sequence of a separate chromogranin. Chromogranin A differs by one amino-acid of its first 14 from the sequence of β -granin, a 21K protein co-secreted with insulin from a rat insulinoma cell line¹⁶. It remains to be seen if this variation is tissue- or species-derived. Settlemann *et al.* have

```

      10          30          50
5' GCTCTGGGATCTGCATCTGGTTGCTGGCGCGTACACGAGACCTGCACCACCAACCCATCC
      70          90          110          130
CCGCGCTGTTGTCGCCGACGTTGCTGGAGCGAGCAGTCCAGCCGCCCTCGCCGAGCGCGCGCG
      150          170          190
TGGCCCCCCCCAGACCACGACTGCTCGGCGCCCGCTTCGCC ATG CGC TCC GCC GCG
                                     met arg ser ala ala
      210          230
GTC CTG GCG CTT CTG CTC TGC GCG GGG CAA GTC ATT GCC CTG CCT GTG AAC
val leu ala leu leu leu cys ala gly gln val ile ala leu pro val asn
      250
AGC CCC ATG AAT AAA GGG GAC ACT GAG GTG ATG
ser pro met asn lys gly asp thr glu val met lys cys ile val glu val
[-----14-mer-----]
[-----30-mer-----]
--- --- ---3'
ile ser asp

```

Fig. 1 Determination of the primary sequence of bovine prechromogranin A mRNA. A, Primer extension/sequencing of the 5' region of chromogranin A mRNA. The sequence shown is that of the 271 nucleotides of the 5' end of the chromogranin A mRNA from within the protein coding region to the putative cap site obtained by primer extension and Maxam-Gilbert sequencing. The signal sequence preceding the N-terminal leucine of chromogranin A is underlined. A 16-fold degenerate oligonucleotide probe (5'-AT(A/G)CA(T/C)TTCAT(A/T/G/C)AC-3') corresponding to the region of common N-terminal sequence of bovine chromogranin A and secretory protein I (ref. 3) shown above was synthesized using phosphoramidite chemistry on a Biosearch SAM polynucleic acid synthesizer. This probe hybridized primarily to a 2,100-base mRNA upon Northern blot analysis of bovine adrenomedullary mRNA. The probe was 5'-labelled with polynucleotide kinase and co-precipitated in ethanol with 10 μ g of polyadenylated bovine adrenomedullary RNA at a molar ratio of 100:1 (assuming a 2 kb mRNA comprising 0.05% of total mRNA). Primer extension yielded a single predominant radioactive band of about 285 nucleotides which was sequenced by the method of Maxam and Gilbert²⁵. This sequence information was used to construct the 30-nucleotide probe shown, on an Applied Biosystems Model 380 A polynucleic acid synthesizer, which was used to screen the bovine adrenomedullary clone bank (Fig. 1B). The 30-mer probe hybridized primarily to a 2,100 base mRNA upon Northern blot analysis of bovine adrenomedullary mRNA as well. The Northern hybridization pattern for both the 14-mer and 30-mer probes was indistinguishable from that obtained with the pCHRG4A cDNA insert described in B, as shown in Fig. 4. B, Restriction analysis and sequencing of pCHRG4A. Polyadenylated RNA was extracted from bovine chromaffin cells with guanidinium hydrochloride, followed by oligo(dT)cellulose chromatography²⁶. A cDNA library was constructed with 4 μ g of poly(A)+RNA and 2 μ g poly(dT)-tailed pCDV1 vector-primer according to the method of Okayama and Berg²⁷. 140 ng of the recombinant plasmid preparation was transfected into competent *E. coli* DH1 cells by the rubidium chloride procedure²⁸ producing 100,000 transformants, which were screened by colony hybridization²⁶ with the 30-nucleotide probe (see Fig. 1) labelled with γ -³²P-ATP and T4 polynucleotide kinase²⁶. The oligonucleotide hybridized to four clones from the library. Restriction analysis suggested that three of the clones were 5' deletions of the fourth and longest clone, pCHRG4A. This clone, containing an 1,883-base insert, was sequenced by the method of Sanger²⁵ after sub-cloning overlapping restriction fragments into M13 mp18 and 19 using JM101 as the bacterial host. Solid arrows show M13 sequencing strategy. Some fragments depicted (dotted arrows) were sequenced the method Maxam and Gilbert²⁵ method.

argued that chromogranin A contains repeating subunits of homologous but non-identical eicosapeptide sequences, based on the partial sequencing of cyanogen bromide fragments of a putative chromogranin A protein¹⁷. We find none of these sequences in pCHRG4A, but rather a single sequence homologous (50-78%) to all of them (Fig. 1B residues 303-326). Neither this nor any other sequence motif (with the exception of oligoglutamic acid) is repeated within the chromogranin molecule. Similar but non-identical sequences found within various chromogranin polypeptides may arise from wholly separate but similar proteins.

Chromogranin A is identical to secretory protein I in the first twenty amino-acids of the mature protein, as well as the residues of the 18 amino-acid signal peptide deduced from radioactive partial sequencing by Majzoub *et al.*¹⁸. Hamilton *et al.*¹⁹ have analysed the amino-acid composition of cyanogen bromide fragments common to chromogranin A and secretory protein I. The largest fragment is identical in composition to the 25K fragment flanked by the fourth and fifth methionine of the protein sequence deduced from pCHRG4A (Fig. 1B, residues 51-277). This together with the detection of chromogranin mRNA as an abundant species in bovine parathyroid (see below) indicates the primary sequences of these two proteins are identical. The

‡ To whom correspondence should be addressed.

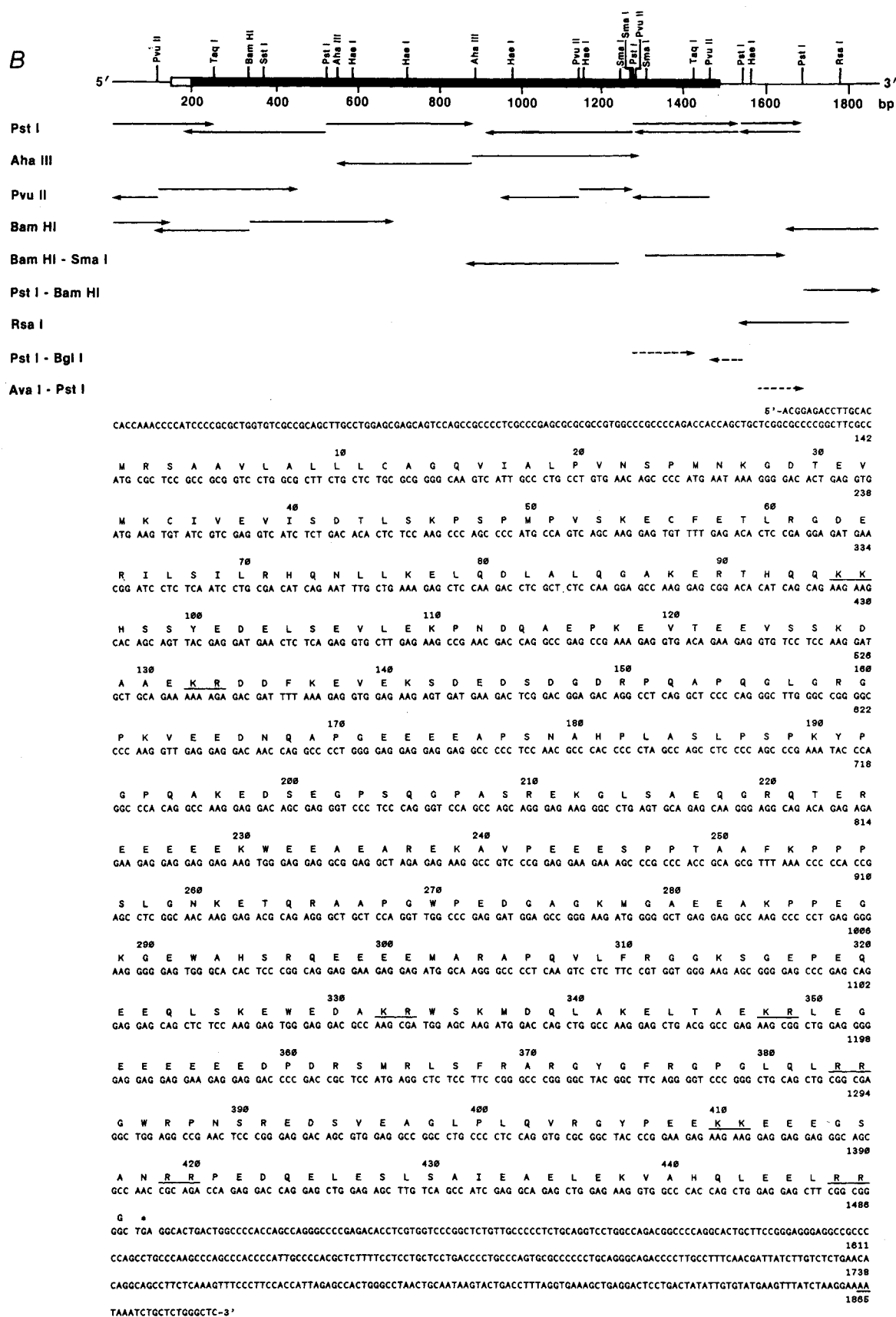


Fig. 1B

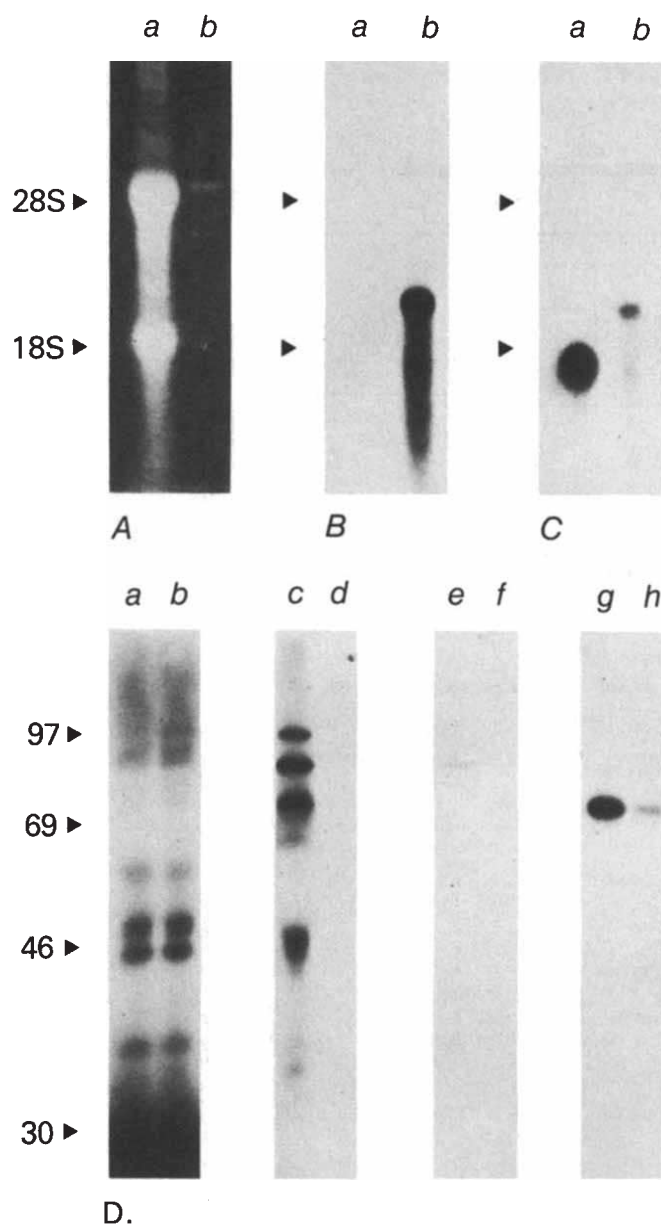


Fig. 2 Immunologic identification of chromogranin A encoded by pCHRG4A. **A**, Ethidium stained gel containing 5% of total input RNA not selected by pCHRG4A_{Xho} (lane *a*) and 25% of total input RNA which was selected by hybridization to pCHRG4A_{Xho} (lane *b*). **B**, Northern blot of gel depicted in **A**, hybridized with ³²P-labelled pCHRG4A. Note that the putative chromogranin cDNA probe hybridizes to a 2,100-base mRNA species only in the selected, and not at all in the nonselected mRNA population. **C**, Northern blot depicted in **B** re-hybridized with ³²P-labelled pbovenk (400 base *Pst*I fragment of the cDNA encoding bovine preproenkephalin A²⁹). Signal in lane *b* is residual radioactivity from the first hybridization with the pCHRG4A probe. Note that the bovine enkephalin cDNA probe hybridizes to the 1,400-base enkephalin mRNA species only in the nonselected mRNA population. **D**, Lanes *a* and *b*, total ³⁵S-labelled protein synthesized under the direction of total polyadenylated chromaffin cell RNA. Lane *a* represents proteins present after immunodepletion of the reticulocyte lysate with antichromogranin A antiserum and lane *b* represents total proteins present in the lysate mixture. Lanes *c* and *d*, immunoprecipitates of cell-free translations of total chromaffin cell polyadenylated RNA, carried out in the presence (lane *d*) or in the absence (lane *c*) of excess unlabelled antigen. Note that the antiserum employed immunoprecipitates three major protein bands, of apparent *M*_s of 75, 87 and 100K (protein size markers are shown on the left). The band at 75K is chromogranin A. Lanes *e* and *f*, immunoprecipitates of cell-free translation of total chromaffin cell RNA not selected by pCHRG4A_{Xho} (A–C, lanes *a*) in the presence (lane *f*) or absence (lane *e*) of excess unlabelled antigen. Lanes *g* and *h*, immunoprecipitates of cell-free translation of total chromaffin cell RNA selected by pCHRG4A_{Xho} (A–C, lanes *b*) in the presence (lane *h*) or absence (lane *g*) of excess unlabelled antigen.

Methods. Hybrid selection: The entire cDNA insert pCHRG4A plus 187 bases of 5' and 30 bases of 3' simian virus 40 flanking sequences was removed from the vector by digestion with *Xho*I, and 2 µg spotted onto nitrocellulose and used to select chromogranin-specific mRNA from 100 µg of total bovine chromaffin cell RNA²⁶. Aliquots of RNA not hybridizing to the cDNA (lane *a* in A–C) and RNA selected by hybridization with the cDNA (lane *b* in A–C) were electrophoresed on a 1% agarose formaldehyde gel, stained with ethidium bromide to visualize 18 and 28S RNA and transferred to Gene-Screen (NEN) paper for sequential hybridization with ³²P-nick-translated pCHRG4A (1,250 base *Hinf*I fragment) or pbovenk at 55 °C in 40% formamide/5 × SSC/50 mM NaPO₄/0.1% polyvinylpyrrolidone/0.1% Ficoll/0.1% bovine serum albumin (BSA)/0.1% SDS/250 µg ml⁻¹ yeast tRNA/250 µg ml⁻¹ denatured herring sperm DNA/10% dextran sulphate. The blot was washed in 0.2 × SSC/0.1% SDS at 55 °C after each hybridization prior to autoradiography on X-Omat AR X-ray film for 20 min. Between hybridization with each probe, the blot was washed in H₂O at 98 °C for 10 min. **Cell-free translation.** 20% of the hybrid-nonselected or the hybrid-selected RNA and one µg of bovine chromaffin cell polyadenylated RNA were used for cell-free translation, carried out using a commercially available rabbit reticulocyte lysate system and ³⁵S-methionine (NEN) according to the manufacturer's recommendations. Each translation mixture was divided in half, and immunoprecipitated in the presence or in the absence of 2 µg of authentic unlabelled bovine chromogranin A purified by one-dimensional electrophoresis (courtesy of Dr Claude Gagnon). Immunoprecipitation was carried out at 4 °C for 24 h in 500 µl buffer (100 mM NaPO₄, pH 7.5, 10 mM EDTA, 0.1% BSA, 0.5 TIU ml⁻¹ Trasylol, 0.01% methionine and 0.1% Triton X-100) containing 5 mM methionine, and 0.2 µl Europium anti-chromogranin A rabbit antiserum (the properties of this antiserum in radioimmunoassay and cell-free translation are described in detail elsewhere³⁰). Five microlitres of goat anti-rabbit gamma globulin antiserum (Calbiochem) in 100 µl of buffer C was added and incubation continued for 24 h at 4 °C. The pellet obtained after centrifugation at 6,000g for 10 min was washed, denatured in 1% SDS, 1% β-mercaptoethanol in 50 mM Tris, pH 6.5, and electrophoresed on a 10% polyacrylamide gel in 25 mM Tris, 190 mM glycine, 0.1% SDS at pH 8.8. Gels were fixed, dried and fluorographed.

differential distribution of chromogranin A and secretory protein I previously reported^{4,5} most likely derives solely from tissue-specific differences in post-translational processing of the same primary translation product.

The identification of the protein encoded by pCHRG4A as chromogranin A was positively established in two ways. pCHRG4A was used to enrich for chromogranin-encoding bovine adrenomedullary mRNAs by hybrid selection, and the selected and non-selected RNAs compared by Northern hybridization and cell-free translation/antichromogranin A immunoprecipitation (Fig. 2). The most abundant protein produced by the selected RNA had a *M*_r of 75,000, and was specifically immunoprecipitated by antichromogranin antiserum (Fig. 2). Thus, prechromogranin A is a 50K protein, encoded by a 2,100-base messenger RNA, which migrates on SDS-PAGE with an apparent *M*_r of 75,000. Further confirmation of the anomalous behaviour of this protein on SDS-PAGE was obtained by cell-free translation of a synthetic prechromogranin

A mRNA, produced by transcription *in vitro* of a pCHRG4A fragment containing the entire protein-coding region of prechromogranin A, subcloned into the vector pGEM2. The protein generated *in vitro*, of actual *M*_r 50,000, migrated on SDS-PAGE with an apparent *M*_r of 75,000 (Fig. 3).

The primary sequence of chromogranin A suggests two possible functions for this protein. First, chromogranin A may be a precursor for smaller polypeptides which function as hormones upon secretion. Chromogranin A contains eight pairs of basic amino acids (Fig. 1B) which are potential processing signals for the generation of smaller polypeptides. A second function that has been suggested for chromogranin A is the binding of intragranular calcium^{8,9}. Chromogranin A is a relatively high-affinity calcium-binding protein with a *K*_D of 54 µM and a binding capacity of approximately 18 molecules of calcium per molecule of protein (from ref. 9 and the *M*_r of chromogranin A reported here). Chromogranins and calcium are co-stored within secretory vesicles of many endocrine tissues²⁰.

Fig. 3 Cell-free translation of mRNA produced by *in vitro* transcription of pCHRG4A. The *NarI*-*XhoI* fragment (nucleotides 128-1,973, see Fig. 1B) of pCHRG4A, containing the entire protein coding region of the cDNA clone, was subcloned into the polylinker of the pGEM2 transcription vector (Promega Biotec), which is flanked by a T7 polymerase promoter at the 5', and an SP6 polymerase promoter (not shown) at the 3' side. The *NarI*-*XhoI* cDNA insert begins 15 bases upstream from the initiating methionine codon (nucleotide 143, Fig. 1B), extends through the protein coding region (143-1,489, Fig. 1B), and ends 483 bases downstream from the stop codon (1,490) at nucleotide position 1,973 of the cDNA (see Fig. 1B). The recombinant plasmid was linearized with *EcoRI* and transcribed with T7 RNA polymerase according to the methods recommended by the supplier (Promega Biotec). Transcription from the T7 promoter produced an 1,899 base mRNA that contains 1,845 bases from the cDNA insert and 54 bases from the polylinker. One microgram of the *in vitro*-synthesized mRNA was denatured and electrophoresed on a 2.2 M formaldehyde-1% agarose gel. The gel was stained with ethidium bromide to visualize the RNA. **A**, Ribosomal RNA markers are shown in lane *a* and the synthesized mRNA produced by *in vitro* transcription is shown in lane *b*. Approximately 400 ng of synthesized mRNA was translated using a rabbit reticulocyte lysate with ³⁵S-methionine (NEN) according to the methods of the supplier. Half of the translation mixture was denatured in 5% βME-1% SDS and electrophoresed on a 7.5% polyacrylamide-0.1% SDS gel, which was fixed, dried and fluorographed (24 h exposure). **B**, Protein markers are shown in lane *a* and the cell-free translation products of the mRNA that was synthesized by *in vitro* transcription of pCHRG4A are shown in lane *b*. The major product is a protein of apparent *M_r* 75K. The actual *M_r* of this protein is 50K as deduced from the coding potential of the 1,347-base single open reading frame (nucleotides 143-1,490 of pCHRG4A) of the 1,899-base *in vitro*-synthesized mRNA species from which it was translated.

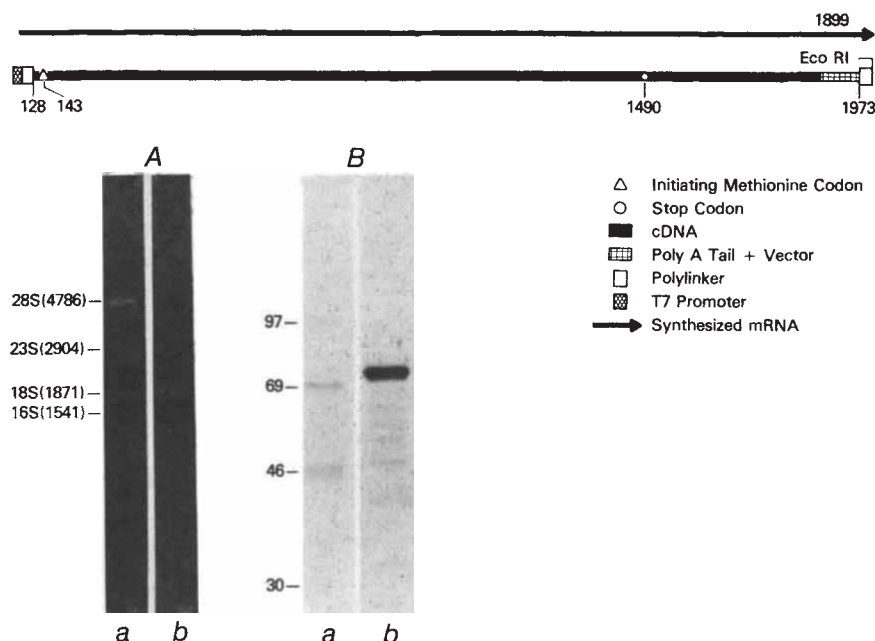


Fig. 4 Northern blot analysis of the tissue distribution of mRNA^{chr8}. Bovine tissue was obtained fresh from a local abattoir, and human and rat tissues immediately after resection. Tissues were frozen on dry ice and stored at -70 °C until used for RNA extraction in guanidinium thiocyanate. Polyadenylated RNA was obtained by oligo(dT)-cellulose chromatography²⁶, suspended in denaturing buffer, electrophoresed on a 1.1% agarose formaldehyde gel, and blotted onto Gene Screen (NEN)(A) or Nytran (Schleicher and Schuell)(B). **A**, 0.1 µg bovine chromaffin cell RNA. Northern blot was pre-hybridized for 3 h at 42 °C, hybridized overnight at 42 °C with a ³²P-labelled nick-translated 1,250-base *HinfI* fragment of pCHRG4A, washed in 0.2 × SSC/0.1% SDS at 42 °C, and the blot autoradiographed for 5 h at -70 °C in cassettes equipped with intensifying screens. Arrows show the positions of migration of *a*, 28S rat ribosomal RNA; *b*, 25S yeast ribosomal RNA; *c*, 23S *E. coli* ribosomal RNA; *d*, 18S rat ribosomal RNA; *e*, 18S yeast ribosomal RNA and *f*, 16S *E. coli* ribosomal RNA. **B**, Bovine adrenal medulla (1), bovine caudate nucleus (2), bovine adrenal medulla (3), bovine anterior pituitary (4), bovine neurointermediate pituitary (5), bovine parathyroid (6), human pheochromocytoma (7), bovine adrenal cortex (8), bovine pancreas (9) and bovine kidney (10). All 1 µg RNA except lane 6, 0.1 µg RNA. Blots were hybridized as described above except that the hybridization was at 60 °C, washing was carried out at 85 °C in 10 mM NaCl/0.1% Na pyrophosphate/1 mM EDTA/0.1% SDS and the probe was an SP6 ³²P-labelled RNA transcript encoding the 5' (antisense) pCHRG4A *PstI*-*PstI* fragment.

Chromogranins may be involved in calcium sequestration and mobilization from secretory vesicles during stimulus-secretion coupling, analogous to the role of calsequestrin during excitation-contraction coupling in muscle cells²¹. Alternatively, calcium binding to chromogranin A during secretory vesicle formation could be a crucial step in organellogenesis in neuroendocrine cells. Chromogranin A has two regions of similarity to the bovine intestinal calcium-binding protein²². These occur in residues 102-113 (7/12 residues identical to CaBP) and residues 323-327 (identical pentapeptide sequence) of the pre-chromogranin A molecule. Chromogranin A also contains a sequence (residues 145-149) identical to a portion of the calcium-binding domain of S100-β²³. S100-β possesses neurotrophic activity *in vitro*²⁴, an activity also ascribed to a low-*M_r* chromogranin obtained from chromaffin granule lysates (K. Unsicker, personal communication).

The pituitary, brain and parathyroid as well as the adrenal medulla are sites of synthesis of prechromogranin A (Fig. 4). Chromogranin A mRNA is absent from non-endocrine tissues such as kidney and endocrine tissues devoid of secretory vesicles such as adrenal cortex. It is present in the rat adrenal medulla (data not shown) and human pheochromocytoma (Fig. 4B), demonstrating considerable conservation of homology among the chromogranins A of these three species. Chromogranin A

mRNA is most abundant in the parathyroid gland, followed by adrenal medulla, pituitary, and brain (Fig. 4). Failure to detect chromogranin A mRNA in whole pancreas may reflect the very small contribution of the endocrine pancreas, in which chromogranin-like immunoreactivity has been observed^{4,5,11,16}, to the total mass of this tissue. Defined probes for chromogranin A mRNA and differentially specific antibodies generated against defined chromogranin protein sequences will provide new tools for the study of neuroendocrine secretion and afford the possibility of *in vivo* diagnosis of pathology throughout the diffuse neuroendocrine system.

We thank Dr Michael Brownstein for providing the oligonucleotide used to screen the adrenomedullary clone bank, Drs Hiroto Okayama and Michael Clarke for advice and assistance in the preparation of the clone bank, Drs Michael Brownstein and Ruth Siegel for their linker and vector primer, Drs Marie-France Bader and Dominique Aunis for the gift of Europium antiserum and Dr Ruth Hogue Angeletti for chromogranin sequence information prior to publication.

Note Added in Proof: While this work was in proof, the sequence of a bovine adrenomedullary cDNA similar, but not identical, to that identified here as the sequence of the prechromograin A mRNA was reported by Benedum *et al.*³¹

Received 28 March; accepted 17 July 1986.

1. Blashko, H., Comline, R. S., Schneider, F. H., Silver, M. & Smith, A. D. *Nature* **215**, 58–59 (1967).
2. Winkler, H. *Neuroscience* **1**, 65–80 (1976).
3. Cohn, D. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6056–6059 (1982).
4. Cohn, D. V., Elting, J. J., Frick, M. & Elde, R. *Endocrinology* **114**, 1963–1974 (1984).
5. O'Connor, D. T., Burton, D. & Deftos, L. J. *Life Sci.* **33**, 1657–1663 (1983).
6. Somogyi, P. *et al. Brain Res. Rev.* **8**, 193–230 (1984).
7. Angeletti, R. H. & Hickey, W. F. *Science* **230**, 89–90 (1985).
8. O'Connor, D. T., Parmer, R. J. & Deftos, L. J. *Trans. Am. Physicians* **97**, 242–250 (1984).
9. Reiffen, F. U. & Gratzl, M. *FEBS Lett.* **195**, 327–330 (1986).
10. O'Connor, D. T. & Bernstein, K. N. *New Engl. J. Med.* **311**, 764–770 (1984).
11. Wilson, B. S. & Lloyd, R. V. *Am. J. Pathol.* **115**, 458–468 (1984).
12. Kiang, W. L., Krusius, T., Finne, J., Margolis, R. U. & Margolis, R. K. *J. biol. Chem.* **257**, 1651–1657 (1982).
13. Kruggel, W., O'Connor, D. T. & Lewis, R. V. *Biochem. biophys. Res. Commun.* **127**, 380–383 (1985).
14. Hogue Angeletti, R. *Arch. Biochem. Biophys.* **184**, 364–372 (1977).
15. Settleman, J., Nolan, J. & Hogue Angeletti, R. *J. biol. Chem.* **260**, 1641–1644 (1985).
16. Hutton, J. C., Hansen, F. & Peshavaria, M. *FEBS Lett.* **188**, 336–340 (1985).
17. Settleman, J., Fonseca, R., Nolan, J. & Hogue Angeletti, R. *J. biol. Chem.* **260**, 1645–1651 (1985).
18. Majzoub, J. A., Kronenberg, H. M., Potts, J. T. Jr, Rich, A. & Habener, J. F. *J. biol. Chem.* **254**, 7449–7455 (1979).
19. Hamilton, J. W., Chu, L. L. H., Rouse, J. B., Reddig, K. & MacGregor, R. R. *Arch. Biochem. Biophys.* **244**, 16–26 (1986).
20. Poisner, A. M. & Trifaro, J. M. (eds) *The Secretory Granule* (Elsevier, Amsterdam, 1982).
21. Ostwald, T. J., MacLennan, D. H. & Dorrington, K. J. *J. biol. Chem.* **249**, 5867–5871 (1974).
22. Wasserman, R. H., Corradino, R. A. & Taylor, A. N. *J. biol. Chem.* **243**, 3978–3986 (1968).
23. Isobe, T. & Okuyama, T. *Eur. J. Biochem.* **89**, 379–388 (1978).
24. Kligman, D. & Marshak, D. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7136–7139 (1985).
25. Davies, R. W. in *Gel Electrophoresis of Nucleic Acids, A Practical Approach* (eds Rickwood, D. & Hames, B. D.) 117–172 (IRL, Oxford, 1982).
26. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
27. Okayama, H. & Berg, P. *Molec. Cell Biol.* **3**, 280–289 (1983).
28. Hanahan, D. in *DNA Cloning, A Practical Approach* Vol. 1 (ed. Glover, D. M.) 109–135 (IRL, Oxford, 1985).
29. Gubler, U., Seeburg, P. H., Hoffman, B. J., Gage, L. P. & Udenfriend, S. *Nature* **295**, 206–208 (1982).
30. Ehrhart, M. *et al. J. Histochem. Cytochem.* **34** (in the press).
31. Benedum *et al. EMBO J.* **5**, 1495 (1986).

Tumour necrosis factor as immunomodulator and mediator of monocyte cytotoxicity induced by itself, γ -interferon and interleukin-1

Ramila Philip & Lois B. Epstein

Cancer Research Institute and Department of Pediatrics, University of California, San Francisco, California 94143, USA

Activated monocytes or macrophages can release soluble cytotoxic molecules capable of lysing tumour cells *in vitro*^{1–6} and thus represent an important component of the host defence mechanisms against malignancy. The recent availability of pure recombinant or natural human lymphokines and monokines and their respective polyclonal or monoclonal antibodies now makes it possible to dissect the interactions of these factors in the induction and performance of the cytotoxic event by the monocytes. Our studies indicate that pretreatment of monocytes with α -IFN or γ -IFN, and also interleukin (IL)-1 or tumour necrosis factor (TNF) results in enhanced monocyte cytotoxicity. Although all these substances induce the production of IL-1 by monocytes, TNF mediates the enhanced cytotoxicity induced in monocytes by γ -IFN, IL-1 and, in an autocrine manner, by TNF itself. Neither TNF, IL-1, γ -IFN nor α -IFN mediate spontaneous monocyte cytotoxicity or that induced by α -IFN. Our studies thus reveal new interactions between the two monokines IL-1 and TNF and provide a dual role for TNF, as immunomodulator and mediator of monocyte cytotoxicity induced by certain specific lymphokine and monokine molecules.

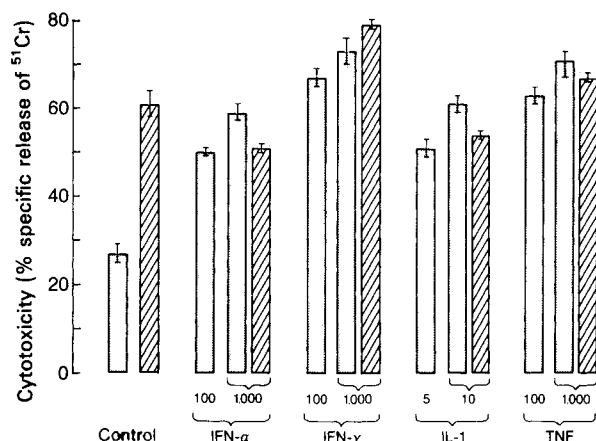


Fig. 1 Enhancement of monocyte cytotoxicity by pretreatment with LPS, α -IFN, γ -IFN and TNF. LPS from *Escherichia coli*, serotype 0127:B8 (Sigma) was used at a final concentration of $1 \mu\text{g ml}^{-1}$. Human recombinant α -IFN was purified to homogeneity and had a specific activity of 4×10^8 IU per mg (a gift from S. Pestka, Roche)³¹. It was used at final concentrations of 100 and 1,000 IU per ml. Human recombinant γ -IFN, purified to homogeneity and had a specific activity of 3.4×10^7 IU per mg (a gift from H. M. Shepard, Genentech)³². It was used at final concentrations of 100 and 1,000 IU per ml. Human natural ultrapure IL-1 with a specific activity of 8×10^6 U μg^{-1} was purchased from Genzyme. This material³³, free of α -IFN, IL-2, and endotoxin, contains both forms of IL-1, with nine parts IL-1 β to 1 part IL-1 α (D. Masters personal communication). It was used in these experiments at final concentrations of 5 and 10 U ml^{-1} . Human recombinant TNF was purified to homogeneity and had a specific activity of 10^8 U mg^{-1} (a gift of H. M. Shepard)³⁴. It was used at final concentrations of 100 and 1,000 U ml^{-1} . The open bars show the result of pretreating the monocytes with each designated agent alone for 18 h before adding them to the tumour target cells; the hatched bars show the result after pretreatment with the agent in the presence of LPS. Each experiment comprises three replicates per group and the results show the mean $\pm$ s.e.m. for seven different experiments.

Methods. Peripheral blood mononuclear cells from normal adult donors were separated by Ficoll-Hypaque gradient centrifugation³⁵ and the monocytes isolated by adherence to plastic. Five ml of RPMI-1640 medium with 10% fetal calf serum (FCS) containing 30×10^6 cells were incubated in 60 mm τ -static petri dishes (Falcon) for 1 h at 37°C in a 5% CO_2 atmosphere. After removing the nonadherent cells by three washes with medium, the adherent cells were removed by incubating them with 0.2% EDTA in phosphate-buffered saline (PBS) with 5% fetal calf serum (FCS) for 20 min. at 37°C . The purity of the monocyte preparations was always $>90\%$ as determined by immunofluorescence microscopy and a fluorescein conjugated antiserum to surface antigen M₃ (Becton Dickinson). Monocytes were then resuspended in RPMI 1640 + 5% FCS and treated either with medium or the designated concentrations of α -IFN, γ -IFN, IL-1, or TNF with or without LPS for 18 h at 37°C at 5% CO_2 in polypropylene tubes (Falcon, 2063). After washing pretreated monocytes were used in a cytotoxicity assay previously described^{7,29} with some minor modifications. Murine fibrosarcoma cells (WEHI 164), a gift from Noel Warner, Becton Dickinson were used as target cells. Before inclusion in the cytotoxicity assay, they were pretreated with $1 \mu\text{g ml}^{-1}$ actinomycin D (Sigma) for 3 h and labelled with ^{51}Cr (NEN; specific activity 200 Ci g) at a concentration of 50 μCi per 10^6 cells for 60 min at 37°C in 5% CO_2 . Effector cells in complete medium were mixed with 5×10^3 targets to give an effector to target ratio of 20:1 in round-bottomed, 96-well microtest plates (Linbro) in a final volume of 0.2 ml per well. The plates were centrifuged at $35 \times g$ for 3 min to enhance effector-target cell contact, and incubated for 8 h at 37°C in 5% CO_2 . The plates were then centrifuged at $350 \times g$ for 6 min and 0.1 ml supernatant collected from each well and counted in a γ -counter. Viability of the monocytes was always $\geq 98\%$ after either the pretreatment phase or the cytotoxicity phase of the assay. Spontaneous release was determined by incubating target cells with complete medium. Maximum release was determined after lysis of the target cells with 100 μl 15% saponin with 2% EDTA. Data were analysed using the formula:

$$\% \text{ specific } ^{51}\text{Cr} \text{ release} = \frac{\text{experimental c.p.m.} - \text{spontaneous c.p.m.}}{\text{maximum c.p.m.} - \text{spontaneous c.p.m.}} \times 100$$

The particular ^{51}Cr -release monocyte cytotoxicity assay selected for our studies (described in the Figure 1 legend) is of short duration and permits rapid assessment of monocyte function before excessive maturation of the cells *in vitro*. Furthermore the target cells used are from a murine fibrosarcoma cell line (WEHI-164) pretreated with actinomycin-D (WEHI-D) that are sensitive only to the cytolytic action of monocytes and not to

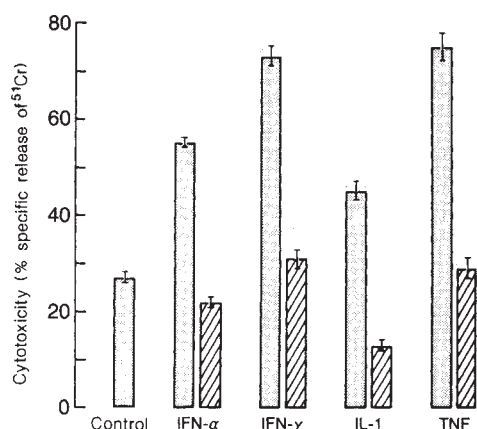


Fig. 2 Pretreatment of monocyte with α -IFN, γ -IFN, IL-1 or TNF and their specific antibodies inhibits enhancement of monocyte cytotoxicity. The source and specific activity of the cytokines used are described in Fig. 1 legend. Polyclonal antibodies to human natural α -IFN were a gift of K. Cantell, Helsinki. Monoclonal antibodies to recombinant human γ -IFN and to human recombinant TNF were a gift from H. M. Shepard (Genentech). Polyclonal antibodies to human natural IL-1 were obtained from Cistron. The open bars reflect the results of pretreating the monocytes with each agent alone for 18 h before adding them to the tumour target cells; the hatched bars show the result after pretreatment with each agent in the presence of its specific neutralizing antibody. The results depicted represent the mean $\pm$ s.e.m. of triplicate samples from each of four experiments.

Methods. Peripheral blood monocytes were isolated as described in the legend to Fig. 1, pretreated for 18 h at 37 °C with the cytokines indicated either with or without their specific antibodies. Various concentrations of antibodies were used but the data depicts the highest concentration. After being washed the monocytes were placed in contact with tumour target cells for 8 h at 37 °C in a short term ^{51}Cr -release assay as described in Fig. 1. Data analysis see Fig. 1 legend.

natural killer cells⁷⁻⁹ (which represent <1% contamination of our monocyte preparations). Using this assay we determined the optimum monocyte: tumour target ratio for maximum cytotoxicity to be 20:1 and used this ratio throughout our studies. We then examined the effects on cytotoxicity of pretreating the monocytes with either endotoxin (LPS) or pure α -IFN, γ -IFN, IL-1, or TNF. Figure 1 shows that treatment with LPS alone results in a greater than two-fold increase in cytotoxicity. Comparable increases are also observed when the monocytes are pretreated with either α -IFN, γ -IFN, IL-1 or TNF in the absence of LPS. When LPS is combined with each of these cytokines no significant further enhancement of cytotoxicity is observed.

It is unlikely that the enhanced monocyte cytotoxicity observed in response to α -IFN, γ -IFN or TNF reflects contamination of any of these reagents with other lymphokines or monokines as they are recombinant products, and the natural IL-1 was stated to be free of α -IFN, interleukin (IL)-2 and endotoxin. Furthermore, the levels of endotoxin detected by the *limulus* amoebocyte lysate (LAL) assay in the γ -IFN were $<0.125 \text{ ng ml}^{-1}$ and in TNF were $<0.8 \text{ ng ml}^{-1}$ which is a level at least 500-fold less than the lowest concentration of LPS capable of stimulating any enhancement of monocyte cytotoxicity (that is, 500 ng ml^{-1}). Further evidence of the ability of each cytokine to enhance monocyte cytotoxicity comes from experiments in which we compared the effects on cytotoxicity of pretreating the monocytes with each of the four agents either alone or in the presence of specific monoclonal or polyclonal antibodies against the agents (Fig. 2). In each experiment the presence of antibody against the cytokine during the pretreatment phase reduced the monocyte cytotoxicity to baseline control levels, and in the case of IL-1, levels somewhat below the

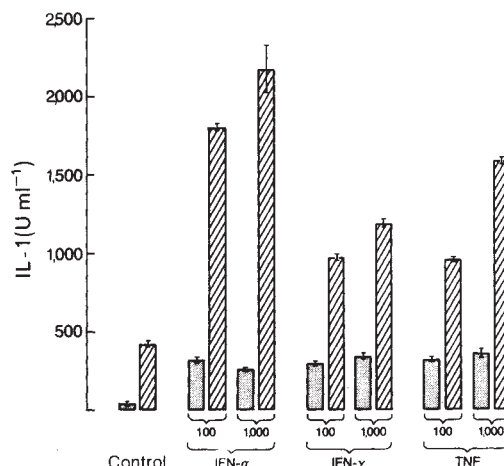


Fig. 3 Induction of monocyte IL-1 by α -IFN, γ -IFN, and TNF and their synergy with LPS. The source and specific activity of the cytokines used in these experiments are described in Fig. 1. The cell line (D10.G4.1) used for the assay of IL-1 was a gift from C. Janeway, Jr, Yale. The open bars represent the results of treating the monocytes for 48 h at 37 °C with the agent at the concentration shown (see Fig. 1 for unit activity of each cytokine); the hatched bars show the result of a treatment with each agent in the presence of LPS. The results are expressed as the mean $\pm$ s.e.m. of duplicate samples from each of seven experiments.

Methods. Peripheral blood monocytes were isolated as described in Fig. 1. Cells (1×10^6) in 0.2 ml medium were cultured in 96-well microtitre plates (Linbro) in the presence of each agent alone or in the presence of LPS, the latter at a concentration of $10 \mu\text{g ml}^{-1}$. After 48 h at 37 °C in 5% CO_2 the culture supernatants were collected and assayed for IL-1 activity. The method measures the incorporation of ^3H -thymidine (specific activity 2 Ci mmol^{-1} , Amersham) into D10.G4.1 cells, a cloned AKR/J T cell line specific for hen-egg conalbumin³⁶. IL-1 units per ml were calculated by deriving the exact dilution at which a sample causes 50% of maximal ^3H -thymidine incorporation, which was determined by linear regression analysis of the maximum c.p.m., minimum c.p.m., and the slope derived from the dilution curve for a standard sample of IL-1.

baseline. In parallel experiments, no abrogation of the enhanced monocyte cytotoxicity in response to α -IFN, γ -IFN, IL-1, or TNF was observed when antibodies other than to the cytokine used for pretreatment were included in the pretreatment phase.

The results of monocyte-mediated cytotoxicity are not always comparable from laboratory to laboratory because of considerable variations in the method of isolation of the monocytes, the choice of target cells, the duration of the pretreatment or cytotoxic phases of the response and the technique used for detecting the extent of cytotoxicity. Nevertheless, several laboratories have observed that LPS and the interferons can enhance monocyte cytotoxicity but in assays of much longer duration (that is, $\geq 72 \text{ h}$)¹⁰⁻¹⁴. Furthermore, with regard to γ -IFN, evidence exists for a specific γ -IFN receptor on cells of the monocyte-macrophage lineage that regulates macrophage tumoricidal activity¹⁵ and which now appears to be quite distinct from the γ -IFN receptor found on non-haematopoietic cells¹⁶.

The observation that IL-1 and TNF, both products of monocytes, can affect monocyte function was particularly intriguing. One other group has also observed enhanced monocyte cytotoxicity in response to IL-1, but again, the assay was of much longer duration than the one we use¹⁷.

We observed that IL-1 and TNF only enhanced monocyte cytotoxicity. In parallel experiments using a short-term ^{51}Cr -release assay with natural killer cells of lymphoid origin as effectors and K562 cells as the tumour-cell targets, we found no enhancement of cytotoxicity after pretreatment of the natural killer cells for 18 h with IL-1 or TNF. In contrast, both α -IFN

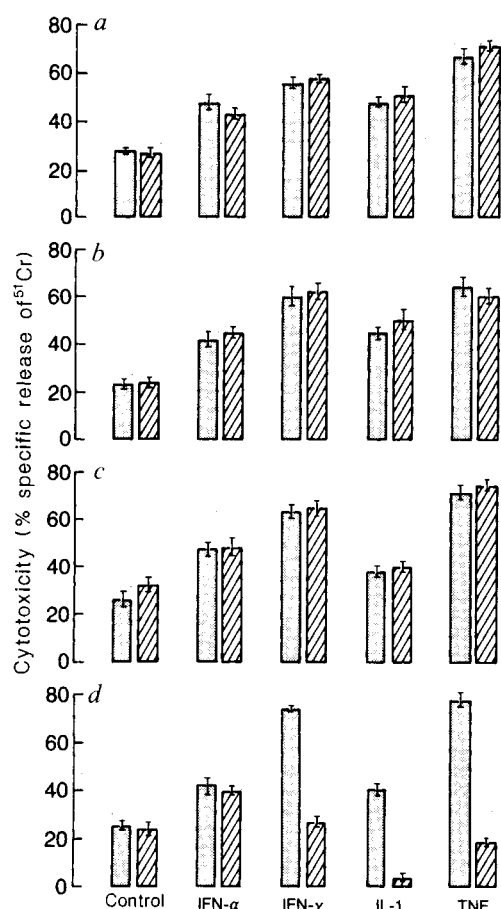


Fig. 4 Effect on monocyte cytotoxicity of antibodies to α -IFN, γ -IFN, IL-1 and TNF added during the cytotoxic phase of the assay. Open bars, agent alone. Hatched bars: a, agent + anti- α -IFN; b, agent + anti- γ -IFN; c, agent + anti-IL-1; d, agent + anti-TNF. The source and specific activity of the cytokines used in these experiments are described in Fig. 1 and the source of the antibodies in Fig. 2. Data are expressed as the mean $\pm$ s.e.m. of triplicate samples from each of three experiments for each antibody used. **Methods.** Peripheral blood monocytes were isolated and treated for 18 h at 37 °C with either α -IFN (1,000 IU ml⁻¹ final concentration), γ -IFN (1,000 IU ml⁻¹), IL-1 (10 U ml⁻¹) or TNF (1,000 U ml⁻¹), as described in Fig. 1. Monocytes pretreated with medium served as controls. After being washed the pretreated monocytes or their controls were combined with the ⁵¹Cr-labelled target cells as described in Fig. 1. Medium was then added to a control group (open bars, and the experimental groups received either specific antibodies to α -IFN (as shown in the top panel), antibodies to γ -IFN (second panel), antibodies to IL-1 (third panel), or antibodies to TNF (bottom panel) in sufficient quantity to neutralize 1,000 IU ml⁻¹ for α -IFN and γ -IFN, 10 U ml⁻¹ for IL-1, and 1,000 U ml⁻¹ for TNF. Read vertically, the effect on monocytes pretreated with α -IFN of antibodies to α -IFN, γ -IFN, IL-1 and TNF added during the cytotoxic phase of the response can be seen. Read horizontally, the effects of a given antibody on monocytes pretreated with each of the four agents or on control, non-pretreated monocytes can be seen. The antibodies were present during the entire 8 h-cytotoxic phase of the assay. At the end of the assay, 0.1 ml supernatant was collected and counted as described in Fig. 1.

and γ -IFN enhanced natural killer cell and monocyte cytotoxicity.

We next wanted to determine the agent responsible for the enhanced monocyte cytotoxicity seen in response to these four cytokines. In view of recent reports that LPS (refs 18, 19), specific antigens²⁰ and the IFNs (refs 21–23) could enhance IL-1 production by cells of the monocyte-macrophage lineage, we con-

sidered this as a possibility. We therefore assayed the supernatants of monocyte cultures stimulated not only with α -IFN and γ -IFN but also with TNF alone or in combination with LPS for the presence of IL-1 (Fig. 3). We found that TNF alone, as well as α -IFN and γ -IFN can enhance IL-1 production. Depending on the concentrations used, the increase in response to α -IFN was 6–7-fold; for γ -IFN, 6–8-fold and for TNF, 7–8-fold. LPS by itself also stimulated IL-1 production. When LPS was combined with α -IFN, γ -IFN or TNF, synergistic induction of IL-1 occurred with all three cytokines with a 4–5 fold increase for α -IFN, 2–3 fold increase for γ -IFN and a 2–4-fold increase for TNF over that observed with LPS alone. Parallel control experiments showed that α -IFN, γ -IFN, and TNF had no effect in the assay for IL-1.

Although IL-1 was produced in response to α -IFN, γ -IFN and TNF, we also wanted to determine if it, or any of the other cytokines, has a direct cytotoxic effect on the tumour target cells. When cultures were prepared without monocytes, the spontaneous lysis of the actinomycin D-treated target cells in the presence of α -IFN was only $4 \pm 1\%$ (s.e.m.); γ -IFN, $8 \pm 2\%$; and IL-1, $7 \pm 1\%$. TNF was the only cytokine to have a direct cytotoxic effect on the tumour cells with $71 \pm 3\%$ release of ⁵¹Cr.

At this point in our investigation TNF became the most likely candidate for the role of mediator of cytokine-enhanced monocyte cytotoxicity. Although α -IFN (refs 24, 25), IL-1 (refs 18, 26), and TNF (refs 27, 28) are known to be products of cells of the monocyte-macrophage lineage, only TNF has a direct cytotoxic effect on the tumour targets used in our assay for monocyte cytotoxicity. Furthermore, a recent report indicates that TNF mediated the enhanced monocyte cytotoxicity induced by LPS (ref. 29). Therefore, to determine whether TNF also mediates the enhanced monocyte cytotoxicity induced by α -IFN, γ -IFN, IL-1 and TNF itself, monocytes were pretreated with all four agents, washed and antibodies specific for each, or for the other agents were added, in the presence of tumour targets during the cytotoxic phase of the response (Fig. 4). No diminution of enhanced monocyte cytotoxicity was observed when monocytes were pretreated with α -IFN, γ -IFN, IL-1 or TNF and then exposed to the tumour cell targets in the presence of antibodies to α -IFN, γ -IFN, or IL-1. In contrast, the presence of antibodies to TNF during the cytotoxic phase of the response caused marked diminution in the cytotoxic capacity of monocytes pretreated with γ -IFN, IL-1 and TNF itself. Thus, TNF can both enhance monocyte cytotoxicity and mediate it in response to γ -IFN, IL-1 and itself. These experiments also indicate that TNF does not mediate the enhanced monocyte cytotoxicity observed with α -IFN. This conclusion is based on the inability of anti-TNF antibody to inhibit the enhancement of cytotoxicity by α -IFN. Although it is possible that α -IFN may have induced IL-1 and the IL-1 in turn could have induced TNF during the course of the experiment, the quantity of anti-TNF antibody used was sufficient to neutralize the maximal amount of TNF that could have been produced. Therefore, some agent other than TNF must be involved in mediating the enhanced cytotoxicity induced by α -IFN. Also, neither TNF, IL-1, γ -IFN nor α -IFN mediate spontaneous cytotoxicity as shown by the lack of effect on spontaneous cytotoxicity when antibodies to each of these agents were present when non-pretreated macrophages were combined with the tumour target cells.

Our observation that TNF is responsible for monocyte cytotoxicity induced by γ -IFN clarifies a previous observation that γ -IFN can induce a membrane bound cytotoxic factor in human monocytes³⁰. Furthermore, G. Wang and D. Goeddel (personal communication) recently showed that γ -IFN can induce TNF in human monocytes and that the induction occurs at the level of the mRNA. Also, our observation that TNF mediates the monocyte cytotoxicity induced by γ -IFN but not α -IFN is consistent with the observation in a murine system that the mechanism of induction of cytotoxic peritoneal

macrophages by γ -IFN differs from that of α -IFN (ref. 14).

The results presented here extend the interactions between the lymphokine, γ -IFN, and the monokines, IL-1 and TNF, and between IL-1 and TNF themselves. It remains to be determined how many other of the wide-ranging effects of γ -IFN and IL-1, in addition to the monocyte cytotoxicity demonstrated here, might also be mediated by TNF.

This work was supported by NIH grants CA27903 and HD17001. We thank Eileen Hunt for typing the manuscript, and Drs Charles Epstein and Hillel Koren for helpful discussions.

Received 18 March; accepted 7 July 1986.

1. Matthews, N. *Immunology* **44**, 135-142 (1981).
2. Cameron, D. J. *J. reticuloendothel. Soc.* **32**, 247-256 (1982).
3. Reidarson, T. H., Levy, W. III, Klostergaard, J. & Granger, G. A. *J. natn. Cancer Inst.* **69**, 879-887 (1982).
4. Adams, D. O. & Nathan, C. F. *Immun. Today* **4**, 166-170 (1983).
5. Sone, S., Tachibana, K., Ishii, K., Ogawara, M. & Tsubura, E. *Cancer Res.* **44**, 646-651 (1984).
6. McKinnon, K. P., Chen, A. R., Argov, S., Love, B. & Koren, H. S. *Immunobiology* **171**, 26-44 (1986).
7. Colotta, F., Peri, G., Villa, A. & Mantovani, A. *J. Immun.* **132**, 936-944 (1984).
8. Ziegler-Heitbrock, H. W. L. & Reithmuller, G. *J. natn. Cancer Inst.* **72**, 23-29 (1984).
9. Koren, H. S. & Herberman, R. B. *J. Leuk. Biol.* **38**, 441-445 (1985).
10. Cameron, D. J. & Churchill, W. H. *J. Immun.* **124**, 708-712 (1980).

11. Peri, G., Polentarutti, N., Sessa, C., Mangioni, C. & Mantovani, A. *Int. J. Cancer* **28**, 143-152 (1981).
12. Herberman, R. B. *et al. Cell Immun.* **67**, 160-167 (1982).
13. Le, J. *et al. J. Immun.* **131**, 2821-2826 (1983).
14. Blasi, E., Herberman, R. B. & Varesio, L. *J. Immun.* **132**, 3226-3228 (1984).
15. Celada, A., Gray, P. W., Rinderknecht, E. & Schreiber, R. D. *J. exp. Med.* **60**, 55-74 (1984).
16. Orchansky, P., Rubenstein, M. & Fischer, D. G. *J. Immun.* **136**, 169-173 (1986).
17. Onozaki, K., Matsushima, K., Kleinerman, E. S., Saito, T. & Oppenheim, J. J. *J. Immun.* **135**, 314-320 (1985).
18. Oppenheim, J. J. & Gery, I. *Immun. Today* **125**, 113-119 (1982).
19. Herman, J., Kew, M. C. & Rabson, A. *Cancer Immun. Immunother.* **16**, 182-185 (1984).
20. Wallis, R. S., Fujiwara, H. & Ellner, J. J. *J. Immun.* **136**, 193-196 (1986).
21. Arenzana-Seisdedos, F., Virelizier, J. L. & Fiers, W. *J. Immun.* **134**, 2444-2448 (1985).
22. Haq, A. U., Rinehart, J. J. & Maca, R. D. *J. Leuk. Biol.* **38**, 735-746 (1985).
23. Candler, R. V., Rouse, B. T. & Moore, R. N. *J. Interferon Res.* **5**, 179-189 (1985).
24. Roberts, N. J., Douglas, R. G., Simons, R. L. & Diamond, M. E. *J. Immun.* **123**, 446-450 (1980).
25. Saksela, E., Virtanen, I., Hovi, T., Secher, D. S. & Cantell, K. *Prog. med. Virol.* **30**, 78-86 (1984).
26. Dinarello, C. A. *J. clin. Immun.* **5**, 287-297 (1985).
27. Carswell, E. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 3666-3670 (1975).
28. Kelker, H. C. *et al. Int. J. Cancer* **36**, 69-73 (1985).
29. Chen, A. R., McKinnon, K. P. & Koren, H. S. *J. Immun.* **135**, 3978-3987 (1985).
30. Kildahl-Andersen, O., Espevik, T. & Nissen-Meyer, J. *Cell. Immun.* **95**, 392-406 (1985).
31. Staehelin, T., Hobbs, D. S., Kung, H.-F., Lai, C. Y. & Pestka, S. *J. biol. Chem.* **256**, 9750-9754 (1981).
32. Gray, P. W. *et al. Nature* **295**, 503-508 (1982).
33. Rosenwasser, L. & Dinarello, C. *Cell Immun.* **63**, 134-142 (1981).
34. Pennica, D. *et al. Nature* **312**, 724-729 (1984).
35. Boyum, A. *Scand. J. clin. Lab. Invest.* **21**, Suppl. 97, 77-89 (1968).
36. Kaye, J. & Janeway, C. A. Jr *Lymphokine Res.* **3**, 175-182 (1984).

Transposon-dependent mutant phenotypes at the *Notch* locus of *Drosophila*

Simon Kidd & Michael W. Young

The Rockefeller University, 1230 York Ave, New York, New York 10021, USA

Many mutations at complex genetic loci in the fruitfly *Drosophila melanogaster* are associated with insertions of transposable elements. At the *Notch* locus, members of one class of insertion-associated mutations, termed *glossy-like*, produce a recessive viable, smooth-eye phenotype with mottled pigmentation¹. Members of a second class, *facet*, produce a recessive viable, rough-eye phenotype with homogeneous pigmentation¹. Both classes of mutations fail to complement *Notch* lethal mutations, so they behave as *Notch* alleles¹. Here we report that each *glossy-like* mutation is associated with an insertion of the same transposable element (*flea*). Each *flea* insertion occurs in the same orientation, but at different locations within intervening sequences of the *Notch* locus. In contrast, each *facet* mutation is associated with insertion of a unique, non-*flea*, transposable element. Insertions producing a *facet* phenotype and insertions causing a *glossy-like* phenotype can break *Notch* intervening sequences at precisely the same location. This suggests that the type of insertion element rather than its position within an affected gene is the primary determinant of the phenotype observed.

Various *facet* and *glossy-like* mutants and their associated phenotypes, including wing nicking, are shown in Table 1. We analysed DNA from each of the mutants by blotting experiments and, with the exception of *fa^{swbBG}* (see Fig. 1 legend), cloned DNA from each mutant. All the mutations are associated with insertions of foreign DNA in the first or second *Notch* intervening sequence (Fig. 1a).

Restriction maps of the insertions are shown in Fig. 1b. The insertions can be divided into two classes. Insertions in *fa^g*, *fa^{g58}*, *fa^{g62}*, *fa^{fx}* and *fa^{swbBG}* form one class, as all involve the same transposable element, *flea*, and each *flea* insertion occurs in the same orientation with respect to the *Notch* locus. Each of these mutations produces a *glossy-like* phenotype. A partial sequence analysis of *flea* indicates that it is a *copia-like* element (*flea* has long direct end repeats and short inverted terminal repeats; ref. 2 and S.K., unpublished).

Three insertions associated with *fa*, *fa³* and *fa^{Ds}*, form a

Table 1 Phenotypes and insertions associated with eye mutants

Mutant	Insertion	Eye phenotype	Wing nicking
<i>fa^g</i>	<i>flea</i>	<i>glossy-like</i>	No
<i>fa^{g62}</i>	<i>flea</i>	<i>glossy-like</i>	No
<i>fa^{g58}</i>	<i>flea</i>	<i>glossy-like</i>	Yes
<i>fa^{fx}</i>	<i>flea</i>	<i>glossy-like</i>	No
<i>fa^{swbBG}</i>	<i>flea</i>	<i>glossy-like</i>	Yes
<i>fa</i>	<i>opus</i>	<i>facet</i>	Yes
<i>fa³</i>	<i>springer</i>	<i>facet</i>	No
<i>fa^{Ds}</i>	<i>hopper</i>	<i>facet</i>	No

second class. In contrast to the *glossy-like* mutations, cloned segments of DNA from these *facet* mutants contain three different non-*flea* insertions (Fig. 1b), but all three insertions seem to be *copia-like* elements (ref. 2 and S.K., unpublished).

All the insertion elements are transcribed from right to left in the restriction maps of Fig. 1b (Fig. 2). Each of the elements, following insertion, would be transcribed in the opposite direction to *Notch*.

Restriction mapping shows that three *glossy-like* and two *facet*-associated insertions affect the same 0.1-kilobase (kb) region of *Notch* intervening sequence 2 (Fig. 1a). To map these insertions more accurately, we sequenced DNA from four mutants, *fa^{g62}*, *fa^{g58}*, *fa* and *fa³*, and compared the sequences with corresponding wild-type *D. melanogaster* DNA (Fig. 3). The four insertions occurred within 25 base pairs (bp) of each other, and for *fa^{g62}*, *fa* and *fa³*, the duplicated insertion target sequences overlap (Fig. 3). All four insertions occur in runs of TATA residues, or, for *fa^{g58}*, in a variation (TATGTA). *fa^{Ds}*, *fa^{fx}* and *fa^g*, which are separated by up to 3 kb from each other, also involve insertions in TATA residues (Fig. 3). The relative clustering of *glossy-like* and *facet* mutations may reflect a preference for AT-rich insertion sites, and as a consequence, the distribution of AT-rich regions within the two largest intervening sequences at *Notch*.

The primary goal of our study was to determine whether the mutant phenotypes reflect the presence of a diverse array of regulatory functions within *Notch* intervening sequences. To resolve this, it is useful to consider the common attributes of the *glossy-like* and *facet* mutations. First, both classes of mutations are associated with *copia-like* transposable-element insertions, and the inserted elements always interrupt intervening sequences in an orientation such that they are transcribed in the opposite direction to *Notch*. With eight mutants in hand,

Fig. 1 Location of transposable element-insertions associated with eye mutations at *Notch*. **a**, Structure of the predominant 10.4-kb RNA transcribed from the *Notch* locus (from ref. 5). Δ , Locations of insertion elements. Three of the mutations, *fa*, *fa*^s and *fa*³, have been mapped previously by genetic recombination and their order and relative positions with respect to the remaining mutations at *Notch* agree with the physical locations of the insertion elements (refs 2, 6; W. Welshons, personal communication). Genetic recombination analysis shows that *fa*^{g62} and *fa*^{swbBG} also map close to *fa*^s (M.W.Y. and M. Baylies, unpublished), so the insertions are probably the cause of the *facet* and *glossy-like* mutant phenotypes. Coordinate, kilobases; zero, the first *EcoRI* site proximal to the In(1)N⁷⁶⁵⁸ breakpoint in *Notch*². Negative numbers go towards the telomere and positive ones to the centromere. **b**, Restriction maps of insertions. Open blocks, insertion sequences. **c**, Code for restriction enzymes. We constructed maps from single and double digestions of cloned DNA with the indicated restriction enzymes. (The exception is *fa*^{swbBG}, from Bowling Green, which was mapped by genomic blots, indicating that it is similar, if not identical, to *fa*^{g62}: *fa*^{swbBG} does not contain the *strawberry* deletion, and should not be confused with *fa*^{swb} (refs 1, 5).) There are unmapped *PvuII* and *PstI* sites in *fa* and unmapped *PvuI*, *PvuII* and *XhoI* sites in *fa*³. *fa*^{Ds} is unique in that it is from *D. simulans*. Restriction mapping and DNA sequencing indicate a close correspondence of the *D. simulans* and *D. melanogaster Notch* loci (S.K. and M. Kelley, unpublished). Bars under maps of *fa*^s, *fa*^{s58}, *fa*, *fa*³ and *fa*^{Ds}, show extents of probes used in RNA hybridizations (Fig. 2). Cross-hybridization experiments show that *fa*^{s58} DNA contains part of the *flea* end-repeat (Figs 1b, 3). The sub *melanogaster* strain Oregon R. The restriction found inserted within a mutant *Drosophila* tro-

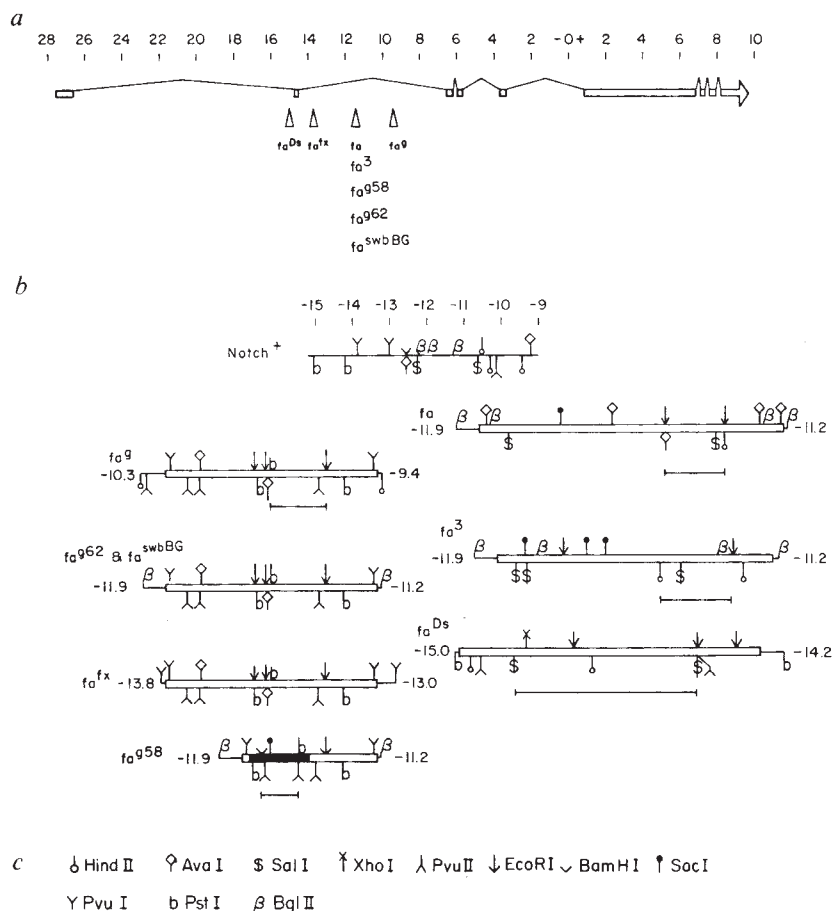


Fig. 2 Direction of transcription of transposable elements. We generated single-stranded RNA probes from the fragments indicated in Fig. 1 after cloning DNA in either pSP64 or pSP65 (ref. 8). ^{32}P -labelled, RNA probes (transcribed from left to right on the insertion element maps of Fig. 1b) were hybridized to 5 μg poly(A) $^{+}$ RNA, isolated as described previously², after the latter was fractionated through a 0.8% agarose, 2.2 M formaldehyde gel and blotted onto Nytran (Schleicher and Schuell). Hybridization conditions were as described elsewhere⁸, except that the formamide and SDS concentrations were 35 and 1%, respectively and the temperature was 65 $^{\circ}\text{C}$. The final wash conditions were $0.1\times\text{SSC}$, 1% SDS at 65 $^{\circ}\text{C}$. For most of the elements, homologous transcripts correspond in size to the insertion sequences and alternate-strand transcription is not clearly detected. An exception is the *springer* insertion of *fa*³. RNA homologous with both DNA strands is detected, but RNA transcribed antiparallel to *Notch* is $\sim 10\times$ as abundantly expressed as RNA transcribed parallel to it. These less-abundant transcripts are much longer (10–11 kb) than the *springer* element (data not shown).

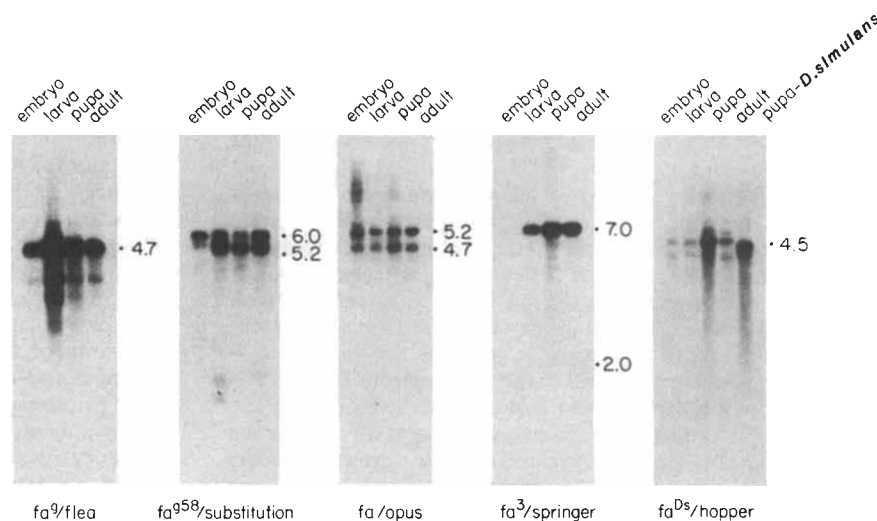


Fig. 3 Sequence analysis of the termini of insertions associated with eye mutants. Left and right ends of each insertion are shown aligned above and below the wild-type sequence, respectively. Sequences from the insertions are in lower-case letters. Duplicated target sequences are shaded (see text). The *fa*^{De} insertion is 127 bp to the right of the *Pst*I site in *D. simulans* at coordinate -15.0. This site is also found in *D. melanocephala*. Restriction-site coordinates are from ref. 2. The *fa*^{fx} insertion is 230 bp to the right of the *Pvu*I site at -13.8, *fa* is 157 bp to the left of the *Bgl*II site at -11.2, and *fa*⁸ is 25 bp to the left of the *Hind*III site at -9.4. All these distances are measured from the 5' nucleotide of the duplicated target sequence to the 5' end of the restriction site. We cloned restriction fragments for sequencing into M13, mp18 or mp19 (ref. 9), and grew and sequenced them as described elsewhere^{10,11}

```

faDe      CCCTTAAAGGCAGCAAAACATACATATActtacgtagttatta
Notch+    CCCTTAAAGGCAGCAAAACATACATATATGTGGGATCTCTCGATCTCTGCAAAC
           atattatgctacATATATGTGGGATCTCTCGATCTCTGCAAAC faDe

fafx      GGATTTTATTGGGTTGTATATATgtaagattgt
Notch+    GGATTTTATTGGGTTGTATATATATAATTTTTTTATATCTGGCATCTGTGTTT
           ccttactccatgttacATATATATAATTTTTTTATATCTGGCATCTGTGTTT fafx

fag58     TTTTCTTATATACGTATGTTATGTATATATATCTCTTCATATATGTATgtaaga
fa       TTTTCTTATATACGTATGTTATGTATATATATAGttccacttgcacagggttctcg
fag       TTTTCTTATATACGTATGTTATGTATATATAaattaataaagtattgtg
fag62     TTTTCTTATATACGTATGTTATGTATATATgtaagattg
Notch+    TTTTCTTATATACGTATGTTATGTATATATATCTCTTCATATATGTATAGGTGT
           gcttactccatgttacATATATATCTCTTCATATATGTATAGGTGT
           ccggttaacttagttaactTATATATATCTCTTCATATATGTATAGGTGT
           aagaagagggttcttaactTATATATCTCTTCATATATGTATAGGTGT
           gcttactccatgttacATGTATAGGTGT fag58

fag       CATGGGAATCCCAAGCAAAAAAAAAAAAAAAAAATATATATATgtaagattgtt
Notch+    CATGGGAATCCCAAGCAAAAAAAAAAAAAAAAAATATATATATATAGATTTAATGAGAG
           cttactccatgttacATATATATAGATTTAATGATAG fag

```

it is reasonable to conclude that deletions, point mutations and even *copia-like* transposable-element insertions with the same transcriptional orientation as *Notch*, do not give eye phenotypes of this sort when they break *Notch* intervening sequences. The phenotypic effect of deleting a portion of an intervening sequence is unknown, but one *copia-like* insertion in this region (in intervening sequence 1), a *B104/roo* insertion associated with a recessive lethal phenotype 1(1)N, is transcribed in the same direction as *Notch* (ref. 2). Also, an FB-transposable element (not a *copia-like* element) insertion has been detected in this intervening sequence in a wild-type strain of *Drosophila*³. Thus, the aberrant eye phenotypes do not seem to result from simply dissociating eye-specific control regions of a *Notch* intron.

Second, the *glossy-like* mutations are found only in conjunction with insertions of members of a single transposable-element family, *flea*. Presumably, a *glossy-like* phenotype cannot be produced by association with any other element. The dependence of the *glossy-like* phenotype on *flea* is also evident from sequencing insertion sites. The insertion sites occupied by *fa* and *fa*^{g62} differ by one nucleotide, whereas the *fa*^{g62} and *fa*³ insertion sites differ by only three nucleotides (Fig. 3), yet *fa* and *fa*³ give a *facet* phenotype and *fa*^{g62} produces a *glossy-like* eye. Clearly the differences in the phenotypes of these mutants reflect special properties of the associated transposable elements rather than where within *Notch* the elements have been integrated.

The insertions associated with *facet* and *glossy-like* mutations probably interfere with *Notch* expression to generate aberrant eye phenotypes because temperature-sensitive, lethal mutations that map to *Notch* protein-coding regions produce similar eye phenotypes following temperature pulses early in pupation^{2,4,5}. Also a short deletion *strawberry*, near the 5' end of *Notch* produces a related mutant phenotype, presumably by altering the level of *Notch* locus transcription⁵. It is possible that insertions associated with *facet* and *glossy-like* mutations only differ in the time or degree to which they interfere with *Notch* activity. Analysis of phenotypes of *fa* and *fa*³ mutants at different temperatures provides support for this notion. At 25 °C both mutants produce a *facet* phenotype, whereas at 29 °C the eyes of these mutants tend toward to *glossy-like* (J. Renvall and S.K., unpublished). The most likely source of any interference with *Notch* locus activity is the expression of the transposable elements themselves, and the elements *flea*, *opus*, *hopper* and *springer* do show different patterns of transcription during *Drosophila* development (Fig. 2). In contrast, analysis of *glossy-like* and *facet* mutations does not suggest a major role for specific sequences within *Notch* introns in the generation of wild-type eye phenotypes. If the interactions of *Notch* and these transposable elements are typical for complex loci in *Drosophila*, it seems likely that the variety of insertion-associated mutant phenotypes

observed for such loci does not accurately reflect the level of informational complexity to be found within corresponding wild-type genes.

We thank Laurel Eckhardt, Peter Model and Norton D. Zinder for comments on the manuscript. We also thank M.M. Green, O. G. Fahmy, P. Portin, W. Welshons and R. Woodruff for supplying *facet* and *glossy-like* mutants, and J. Renvall for assistance in cloning some of the mutants. This work was supported by grants to M.W.Y. from the National Institutes of Health and the Andre and Bella Meyer Foundation.

Received 30 April; accepted 19 June 1986.

1. Lindsley, D. L. & Zimm, G. *Drosophila Information Service* **62**, 114-116 (1985).
2. Kidd, S., Lockett, T. J. & Young, M. W. *Cell* **34**, 421-433 (1983).
3. Grimwade, B. G., Muskavitch, M. A. T., Welshons, W. J., Yedovnick, B. & Artavanis-Tsakonas, S. *Dev Biol.* **107**, 503-519 (1985).
4. Shellenbarger, D. L. & Mohler, J. D. *Dev Biol.* **62**, 432-446 (1978).
5. Kidd, S., Kelley, M. R. & Young, M. W. *Molec. Cell Biol.* **6**, 3094-3108 (1986).
6. Welshons, W. J. *Genetics* **76**, 775-794 (1974).
7. Karlik, C. C. & Fyrberg, E. A. *Cell* **41**, 57-66 (1985).
8. Melton, D. A. et al. *Nucleic Acids Res.* **12**, 7035-7056 (1984).
9. Norrander, J., Kempe, T. & Messing, J. *Gene* **26**, 101-106 (1983).
10. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
11. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963-3965 (1983).

Corrigendum

DNA fingerprint analysis in immigration test-cases

W. G. Hill

Nature **322**, 290-291 (1986).

IN this Matters Arising item, the equation in the footnote to Table 1 should read:

$$\ln L = \sum_i \sum_j \sum_k n_{ijk} [\ln p_{ijk} - \ln (1 - p_{000})]$$

In the calculations as published a correct version of the formula was used.

Errata

Increased levels of myelin basic protein transcripts in virus-induced demyelination

K. Kristensson, K. V. Holmes, C. S. Duchala, N. K. Zeller, R. A. Lazzarini & M. Dubois-Dalq

Nature **322**, 544-547 (1986)

IN the title of this letter the word 'gene' was incorrectly included. The title is correct as above.

Computer Science

Charles Babbage Institute Reprint Series for the History of Computing, Vol. 9, the Moore School Lectures: Theory and Techniques for Design of Electronic Digital Computers. MARTIN CAMPBELL-KELLY and MICHAEL R. WILLIAMS (eds). *The MIT Press/Tomash Publishers*:1985. Pp.568. Hbk ISBN 0-262-03109-4. \$50.

Dictionary of Computing, 2nd Edn. *Oxford University Press*:1986. Pp.416. Hbk ISBN 0-19-853913-4. £17.50.

Expert Systems 85: Proceedings of the Fifth Technical Conference of the British Computer Society Specialist Group on Expert Systems. MARTIN MERRY (ed.). *Cambridge University Press*:1985. Pp.334. ISBN 0-521-32596-X. £22.50.

IBM's Early Computers. By CHARLES J. BASHE, LYLE R. JOHNSON, JOHN H. PALMER and EMERSON W. PUGFH. *The MIT Press*:1986. Pp.716. Hbk ISBN 0-262-02225-7. Np.

A Pascal Database Book. By JULIAN ULLMANN. *Oxford University Press*:1986. Pp.332. Pbk ISBN 0-19-859642-1. Hbk £25; pbk £12.50.

Probabilistic Properties of Deterministic Systems. By A. LASOTA and M.M. MACKEY. *Cambridge University Press*:1986. Pp.358. ISBN 0-521-30248-X. £40, \$49.50.

Program Evolution: Processes of Software Change. M.M. LEHMAN and L.A. BELADY (eds). *Academic*:1985. Pp.538. Hbk ISBN 0-442440-6. £35, \$40.

Chemistry

Advances in Inorganic Chemistry and Radiochemistry, Vol.29. H.J. EMELEUS and A.G. SHARPE (eds). *Academic*:1985. Pp.335. Hbk ISBN 0-12-023629-X. £56.50, \$65.

Asymmetric Synthesis, Vol.5, Chiral Catalysis. JAMES D. MORRISON (ed.). *Academic*:1985. Pp.391. Hbk ISBN 0-12-507705-X. £73.50, \$85.

Chemistry of Organo-Zirconium and -Hafnium Compounds. By D.L. CARDIN, M.F. LAPPERT and C.L. RASTON. *Ellis Horwood*:1986. Pp.451. Hbk ISBN 0-85312-475-2. £59.50.

Chemistry of the Monoterpenes: An Encyclopedic Handbook, Part A & B. By WILLIAM F. ERMAN. *Dekker*:1985. Pp.1709. (Part A) Hbk ISBN 0-8247-1573-X, (Part B) Hbk ISBN 0-8247-7312-8. (USA & Canada) \$145, (other countries) \$174.

Concise Dictionary of Chemistry. *Oxford University Press*:1986. Pp.304. Hbk ISBN 0-19-866143-6. £7.95.

Flavor Chemistry of Fats and Oils. DAVID B. MIN and THOMAS H. SMOUSE (eds). *American Oil Chemists Society*:1985. Pp.309. Hbk ISBN 0-935315-12-8. \$55.

History of Polyolefins. RAYMOND B. SEYMOUR and TAI CHENG (eds). *D. Reidel*:1985. Pp.336. Hbk ISBN 90-277-2128-9. Dfl135, £37.50, \$49.50.

Microdomains in Polymer Solutions. PAUL DUBIN (ed.). *Plenum*:1985. Pp.457. Hbk ISBN 0-306-42110-0. Np.

Modern Aspects of Electrochemistry, No.10. B.E. CONWAY, RALPH E. WHITE and J.O.M. BOCKRIS (eds). *Plenum*:1985. Pp.512. Hbk ISBN 0-306-42024-4. Np.

Modern Bioelectrochemistry. FELIX GUTMANN and HENDRIK KEYZER (eds). *Plenum*:1986. Pp.627. Hbk ISBN 0-306-41981-5. Np.

Opportunities in Chemistry. Committee to Survey Opportunities in the Chemical Sciences, Board on Chemical Sciences and Technology, Commission on Physical Sciences, Mathematics and Resources and National Research Council. *National Academy Press*:1985. Pp.344. Pbk ISBN 0-309-03594-5. £18.

Petrochemicals. By PETER WISEMAN. *Ellis Horwood/Wiley*:1986. Pp.182. Hbk ISBN 0-85312-741-7. £19.50.

Photochemistry, Vol. 16. D. BRYCE-SMITH (ed.). *Royal Society of Chemistry*:1985. Pp.592. ISBN 0-85186-145-8. £125, \$225.

Polymer Yearbook 2, 2nd Ed. RICHARD A. PETHERICK (ed.). *Harwood*:1985. Pp.414. Hbk ISBN 3-7186-0274-1; pbk ISBN 3-7186-0276-8. Hbk \$50; pbk \$25.

Scaling Phenomena in Disordered Systems. ROGER PYNN and ARNE SKJELTORP (eds). *Plenum*:1985. Pp.580.

Selected Developments in Catalysis. J.R. JENNINGS (ed.). *Blackwell Scientific*:1985. Pp.162. ISBN 0-632-01454-7. £37.

Supported Metal Complexes. By F.R. HARTLEY. D. Reidel:1985. Pp.318. Hbk ISBN 90-277-1855-5. Dfl.165, £45.75, \$59.

Earth Sciences

Alkaline Ring Complexes in Africa. *Journal of African Earth Sciences*, Vol. No.1/2 1985. R. BLACK and P. BOWDEN (eds). *Pergamon Press*:1985. Pp.281. Pbk ISSN 0731-7247. \$46.75.

Coastal Wetlands. HAROLD H. PRINCE and FRANK M. D'ITRI (eds). *Lewis/Wiley*:1985. Pp.286. ISBN 0-87371-052-5. £36.75.

Coring Operations: Procedures for Sampling and Analysis of Bottomhole and Sidewall Cores. Written and Compiled by EXLOG staff. D. Reidel:1985. Pp.174. Hbk ISBN 90-277-1980-2. Dfl.115, £31.95.

Dentrification in the Nitrogen Cycle. HAN L. GOLTERMAN (ed.). *Plenum*:1985. Pp.295. Hbk ISBN 0-306-42104-6. Np.

Designing Optimal Strategies for Mineral Exploration. Br J.G. DE GEOFFROY and T.K. WIGNALL. *Plenum*:1985. Pp.364. Hbk ISBN 0-306-41977-7. Np.

Earth Surface Systems. By RICHARD J. HUGGETT. *Springer-Verlag*:1985. Pp.270. ISBN 0-387-15421-3. DM148.

Field Geologist's Training Guide. Written and compiled by EXLOG staff. D. Reidel:1985. Pp.291. Hbk ISBN 90-277-1932-2. Dfl135, £35.50.

Geological Structure of Mac. Robertson Land (East Antarctica). By M.G. RAVICH, D.S. SOLOV'EV and L.V. FEDOROV. A.A. Balkema:1985. Pp.354. Hbk ISBN 90-6191-439-6. £20.75.

Graphite Intercalation Compounds: Progress of Research in Japan. S. TANUMA and H. KAMIMURA (eds). *World Scientific Publishing*:1985. Pp.337. Hbk ISBN 9971-978-96-2. £35.75.

Interregional Air Pollution Modelling: The State of the Art. S. ZWERVER and J. VAN HAM (eds). *Plenum*:1985. Pp.278. Hbk ISBN 0-306-42108-9. Np.

Introduction to Potential Theory: Fundamental Mathematical and Physical Topics for the Study of Physical Geodesy. By RUDOLPH SIGL. *Abacus Press in association with Herbert Wichmann Verlag*:1985. Pp.229. Pbk ISBN 0-856626-451-2. £12.75, \$20.

Issues in Atmospheric and Oceanic Modelling. Part B, Weather Dynamics. SYUKURO MANABE (ed.). *Academic*:1985. Pp.432. Hbk ISBN 0-12-018849-X. £58.50, \$66.

Landslides: Causes, Consequences and Environment. By MICHAEL J. CROZIER. *Croom Helm*:1986. Pp.252. Hbk ISBN 0-7099-0790-7. £25.

New Sites for Old: A Student's Guide to the Geology of the East Mendips. K.L. DUFF, A.P. HICKEY and M.J. HARLEY (eds). *Nature Conservancy Council*:1986. Pp.189. Pbk ISBN 0-86139-319-8. Np.

Oh! Pascal! By DOUG COOPER and MICHAEL CLANCY. *Norton*:1985. Pp.607. Pbk ISBN 0-393-95515-X. £12.95.

The Permafrost Environment. By STUART A. HARRIS. *Croom Helm*:1986. Pp.276. Hbk ISBN 0-7099-3713-X. £22.95.

Radiolociology. V.V. BOGORODSKY, C.R. BENTLEY and P.E. GUDMANDSEN. D. Reidel:1985. Pp.254. Hbk ISBN 90-277-1983-8. Dfl.160, \$44.50, \$54.

Rates of Chemical Weathering of Rocks and Minerals. STEVEN M. COLMAN and DAVID P. DETHIER (eds). *Academic*:1986. Pp.603. ISBN 0-12-181490-4. \$95, £79.

Reefs and Banks of the Northwestern Gulf of Mexico: Their Geological, Biological and Physical Dynamics. By RICHARD REZAK, THOMAS J. BRIGHT and DAVID W. MCGRAIL. *Wiley*:1985. Pp.259. ISBN 0-471-89379-X. £54.25.

Sandstone Depositional Models for Exploration for Fossil Fuels, 3rd Edn. By GEORGE DE VRIES KLEIN. D. Reidel:1985. Pp.209. Hbk ISBN 90-277-2064-9. Dfl.195, £54.

Storm Depositional Systems: Dynamic Stratigraphy in Modern and Ancient Shallow-Marine Sequences. By THOMAS AIGNER. *Springer-Verlag*:1985. Pp.174. Pbk ISBN 3-540-15321-8. DM 32

Temporal Dynamics of an Estuary: San Francisco Bay. J.E. CLOERN and F.H. NICHOLS (eds). *Dr W. Junk*:1985. Pp.237. Hbk ISBN 90-6193-538-5. Dfl175, £48.50, \$67.

White Waters and Black. By GORDON MACCREAGH. *The University of Chicago Press*:1985. Pp.335. Pbk ISBN 0-226-50016-0. £10.25.

Environmental Science

Air Quality. By THAD GODISH. *Lewis/Wiley*:1985. Pp.372. ISBN 0-87371-019-3. £35.90.

The Biology and Management of the River Dee. DAVID JENKINS (ed.). *Institute of Terrestrial Ecology, Monks Wood Experimental Station, Abbots Ripton, Huntingdon*:1985. Pp.160. Pbk ISBN 0904282-88-0. £6.

Environmental Priorities in India and Sustainable Development. By T.N. KHOSHOO. *Indian Science Congress Association, New Delhi*:1986. Pbk. Np.

Far From Paradise: The Story of Man's Impact on the Environment. By HERBERT GIRARDET and JOHN SEYMOUR. *British Broadcasting Corporation*:1986. Pp.216. Hbk ISBN 0-563-20357-9. £9.50.

Indoor Air and Human Wealth. R.B. GAMMAGE and S.V. KAYE (eds). *Lewis*:1985. Pp.430. Hbk ISBN 0-87371-006-1. £35.90.

Introduction to Environmental Science. By JOSEPH M. MORAN, MICHAEL D. MORGAN and JAMES H. WIERSMA. W.H. FREEMAN:1985. Pp.709. Hbk ISBN 0-7167-1684-4. \$29.95.

Large Scale Water Transfers: Emerging Environmental and Social Experiences. GENADY N. GOLUBEV and ASIT K. BISWAS (eds). *United Nations Environmental Programme (Tycooly Publishing, UK)*:1985. Pp.158. Pbk ISBN 1-85148-006-5. Np.

Principles of Nuclear Radiation Detection: Laboratory Manual. By GEOFFREY G. EICHOLZ and JOHN W. POSTON. *Ann Arbor Science*:1985. Pp.181. Pbk ISBN 0-250-40264-5. £15.60.

Strategies for River Basin Management: Environmental Integration of Land and Water in a River Basin. JAN LUNDQVIST, ULRIK LOHM and MALIN FALKENMARK (eds). D. Reidel:1985. Pp.346. Hbk ISBN 90-277-2111-4. Dfl160, £44.50, \$56.

Applied Biological Sciences

Basic and Applied Mutagenesis, With Special Reference to Agricultural Chemicals in Developing Countries. AMIR MUHAMMED and R.C. von BORSTEL (eds). *Plenum*:1985. Pp.441. Hbk ISBN 0-306-42010-4. Np.

Biofeedback and Sports Science. JACK H. SANDWEISS and STEVEN L. WOLF (eds). *Plenum*:1985. Hbk ISBN 0-306-41995-5. Np.

The Detection of New Drug Adverse Reactions. By M.D.B. STEPHENS. *Macmillan Stockton Press*:1985. Pp.262. Hbk ISBN 0-943818-33-8. £40.

Fifth Cooley's Anemia Symposium. ARTHUR BANK, W. FRENCH and EDWARD C. ZAINO (eds). *New York Academy of Sciences*:1985. Pp.471. Pbk ISBN 0-89766-285-7. Np.

Nonfermentative Gram-Negative Rods: Laboratory Identification and Clinical Aspects. GERALD L. GILARDI (ed.). *Marcel Dekker*:1985. Pp.376. ISBN 0-8247-7370-5. \$75.

Recent Advances in Aquaculture, Vol. 2. JAMES F. MUIR and RONALD J. ROBERTS (eds). *Croom Helm*:1985. Pp.282. ISBN 0-7099-3526-9. £35.

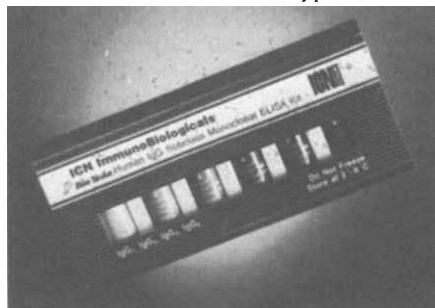
Psychology

Psychiatry: The State of the Art. Vol. 6 Drug Dependence and Alcoholism, Forensic Psychiatry, Military Psychiatry. P. PICHOT, P. BERNER, R. WOLF and K. THAU (eds). *Plenum*:1985. Pp.594. Hbk ISBN 0-306-41606-9. Np.

From cell motility to flow cytometry

This week a scattering of research products encompasses both the unusual and the pragmatic.

In what may be a coup for the company, ICN ImmunoBiologicals has delivered a **monoclonal ELISA kit** that types the four

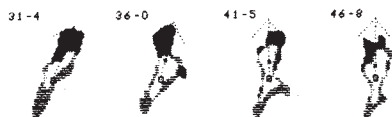


Fine distinctions from monoclonal methods.

subclasses of human IgG (Reader Service No. 100). Intended for research use only, ICN's \$395 (US) kit could give a boost to studies of immunodeficiency diseases, chronic respiratory infections and allergic reactions. The kit uses control serum calibrated to World Health Organization reference preparation 67/97, and provides results in one day.

A digital **delay/pulse generator** with resolution to 5 million millionths of a second can tackle even the most demanding applications, according to Lambda Photometrics (Reader Service No. 101). The DG535 generator has the sort of accuracy normally associated with a laser; each of the DG535's four delay channels has a 1000 s range and a 5 ps resolution. With an internal trigger that operates in either single-shot or burst modes, the generator is programmable from 0.001 Hz to 1 MHz. Lambda sells its generator for £2,550 in the United Kingdom.

If manual methods for tracking cell movement are getting a bit tiresome, the CellTrak system from Motion Analysis could offer a way out (Reader Service No.



Shaped dynamics, described by CellTrak

102). The company coupled IBM PC/AT capabilities with its VP110 video processor to create a **motility analysis package** that, it says, can yield results in 10 to 20 minutes. Automated from data capture to display, CellTrak analysis has the flexibility to describe motion in terms of speed, direction, angular velocities and other parameters. The company's system, priced at approximately \$35,000 (US), can follow both bacterial and mammalian cell movement.

Ortho Diagnostic is handing out **advice on flow cytometry** via its new Applications Support Laboratory (Reader Service No. 103). Ortho's service includes a free telephone consultation line to answer questions on the method, its reagents and applications. Enrolling in the laboratory also nets the user a set of 28 protocols for flow cytometry, with privileges to whatever future applications are published in the series.

Omni Biochem has just released a catalogue that details its manufacturing services and 2,000 **products for biotechnology** and genetic engineering research (Reader Service No. 104). Among the catalogue's pages are descriptions of Omni's custom peptide synthesis service and a variety of biologically active peptides, enzyme substrates and inhibitors and L- and D-amino

**OMNI
BIOCHEM**

Biologically Active Peptides
Omni-Beads® Functionalized Polymers
Amino Acids
Special Amino Acids and Intermediates
Enzyme Substrates and Inhibitors

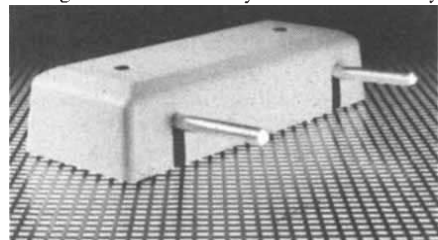


A catalogue airs the wares of Omni Biochem.

acids and their derivatives. Omni also offers more than 100 special amino acids, resins and purified reagents and solvents.

A new process for preparing microorganisms could make **reconstituting bacterial cultures** as easy as mixing instant oatmeal (Reader Service No. 105). The American Type Culture Collection has just begun selling freeze-dried Uniplus cultures that come ready-made with the appropriate growth medium. All that's required to establish a growing broth culture is sterile water and incubation; subculturing is not necessary. Each Uniplus culture, which has a reconstituted volume of 5 ml, costs \$12 (US).

Able to process up to 7.5 l of water per minute, the Fistream UV sterilizer keeps **water purification** rolling (Reader Service No. 106). The UV source Fistream uses in its sterilizer will treat water coming straight from the mains supply or out of storage. Fistream says this flexibility



UV disinfection without delay.

makes the sterilizer ideal for removing microorganisms from a recirculating system or for ensuring sterility after distillation or deionization. □

These notes are compiled by Karen Wright from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENTS

LIPID PRODUCTS

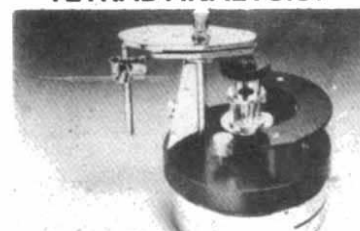
top quality lipids for biomedical research
17 years experience supplying lipids to top research laboratories throughout the world

For free catalogue contact:
Lipid Products, Nutfield Nurseries
South Nutfield, Redhill
Surrey RH1 5PG, UK
Tel: (073 782) 3277

Reader Service No.19

SINGER INSTRUMENTS MICROMANIPULATORS & ACCESSORIES

YEAST? TETRAD ANALYSIS?



Singer MK III Micromanipulator
and Precision Elevator.

Write direct to:
Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset TA23 0QL
England. Tel. (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50

RAW SERUM SPECIALISTS

High Quality
Serum for Cell Culture
— Vaccine Production —
Biotechnological Applications

- * Foetal Calf Serum
- * Newborn Calf Serum
- * Donor Calf and Horse Serum
- * U.K.-New Zealand-Australian Sources



**MEDICAL & VETERINARY
SUPPLIES LIMITED**

Botolph Claydon,
Buckingham MK18 2LR
Tel No: (029 671) 4555

Reader Service No.9

Occupational trends in Britain to 1990

from Richard Pearson

The United Kingdom's biggest employer-based study highlights future employment change.

THE labour market in the 1980s is undergoing a period of prolonged and difficult change throughout the world. Oil price rises, new technologies and increased competition have led to new working patterns and rising levels of unemployment. The recession of 1979–82 has had the most visible effect, pushing unemployment levels higher than ever before, while the collapse in the oil price over the past year has not led to the expected boom in world output and improved employment prospects. As the old industries such as shipbuilding are contracting in the West, and the new technologies are being introduced at an increasingly rapid rate, what will happen to employment patterns over the rest of the decade and what are the implications for occupational change? The results of an employer-based study, cover-

700 jobs, most of them unskilled operatives. In the same year, it opened a new factory, producing a higher volume of a better product using the latest technology, but requiring only 100 staff, more of whom were skilled.

In order to concentrate on their core activities companies are also subcontracting services such as catering, cleaning and legal services to specialist, often smaller companies. As a result, although the production industries are expected to shed around 650,000 jobs over the period to 1990, up to half of them will reappear in the service sector. Thus the apparent decline in production and manufacturing and the growth in services is partly a statistical trick of changing classification.

Nevertheless, the service sector will continue to increase its employment, in part due to this subcontracting, but also due to market growth in sectors such as distribution, financial and leisure services. Not all service activities will grow. Jobs in transport and communication are expected to fall due to a reduction in overcapacity and increased competition, while that in the public sector is expected to fall due to a combination of expenditure restraint, improving working methods, reduced service provision and the transfer of activities to the private sector through privatization and subcontracting.

Within these sectoral shifts, other major changes will be the growing share of employment taken by small companies and the self-employed, and by part-time jobs. Part-time work is expected to account for nearly one in four jobs by 1990. The corollary of this is that full-time "male" jobs will continue to contract, perhaps by as much as a million jobs.

Despite the overall contraction in production-related employment, the employment of engineers, scientists and technologists and also of multi-functional craftsmen and technicians will continue to expand as a result of the growing introduction of new technologies, lessening job demarcations and the increasing emphasis on "value added". A key skill needed by these people, relates to the use and application of information technology, not just in products but also in support activities such as design and marketing. Other important skills for the future will be an ability to work across disciplines, and to develop improved project management and human relations skills. Individuals with these skills will be scarce for sometime.

Other expanding occupations during

this period will be professional staff in areas such as accountancy and marketing as companies seek to improve financial management and increase their understanding of and links with the market and the customer. These skills are seen as increasingly important not just in the private sector but also in the public sector, in both the mainstream services of health and education as well as in administrative activities. Medical and welfare jobs are also expected to continue growing.

Personal and support staff occupations will expand significantly as the service sector expands, resulting in more jobs, main-

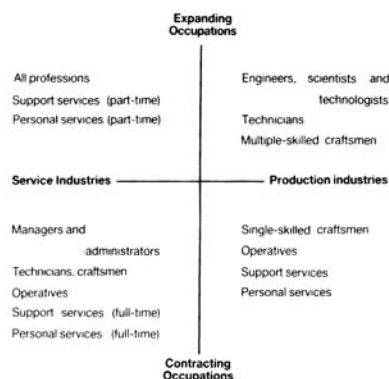


Fig. 1 Predicted sectoral changes in the UK workforce between 1985 and 1990. From the report *UK Occupation and Employment Trends to 1990*, by A. Rajan and R. Pearson (Butterworth, 1986).

ing 3,500 organizations, highlight the key trends for the period to 1990. While total employment levels will be broadly static, there will be major compositional changes.

Output is expected to continue to grow, but the decline in employment in the production industries is likely to continue due to continued weak demand in the United Kingdom and insufficient international competitiveness (Fig. 1). In order to improve competitiveness, employers are expecting to continue to close uneconomic factories, cut back on surplus labour, improve working methods and, particularly in the larger companies, to apply the new technologies to both production and support activities. These developments were illustrated by the food division of a major national company which closed one uneconomic factory with out-dated technology and a low quality product, shedding

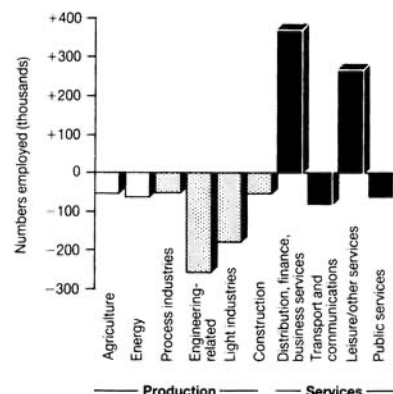


Fig. 2 The changing occupational balance by part-time, in activities such as selling, financial services and catering. Here, staff interacting with customers are being expected to develop broader based social, product diagnostic and entrepreneurial skills to further their organization's success. Occupations in decline will be the predominantly unskilled or single skilled. Across the economy there will be a shift towards occupations with high knowledge content and/or multiple skills (Fig. 2).

In all types of organization, employers see a need to improve and broaden the skills of their workforce and increase their flexibility. This applies particularly to management and professional staff who are viewed as a key to business success in a changing market environment. The challenge will be whether existing management and their workforces can, in conjunction with the external education and training system, increase the level and breadth of skills in the United Kingdom fast enough not just to maintain employment levels but also to increase them and so to reduce unemployment. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.